IRE Transactions on ANTENNAS and PROPAGATION

Volume AP-4 **APRIL 1956** **Number 2**

Published Quarterly

TABLE OF CONTENTS

PUBLISHED BY THE

Professional Group on Antennas and Propagation

IRE TRANSACTIONS PGAP IS A QUARTERLY PUBLICATION DEVOTED TO EXPERIMENTAL AND THEORETICAL PAPERS ON ANTENNAS AND WIRELESS PROPAGATION OF ELECTROMAGNETIC WAVES

MANUSCRIPTS should be submitted to John B. Smyth, Editor, SRA, 3930 4th Avenue, San Diego 3, California. Manuscripts should be original typewritten copy, double spaced, plus one carbon copy. References should appear as footnotes and include author's name, title, journal, volume, initial and final page numbers, and date. Each paper must have an abstract of not more than 200 words. News items concerning PGAP members and group activities should be sent to the News Editor, Mr. H. A. Finke, Polytechnic Research and Development Company, 55 Johnson Street, Brooklyn, New York.

ILLUSTRATIONS should be submitted as follows: All line drawings (graphs, charts, block diagrams, cutaways, etc.) should be inked uniformly and ready for reproduction. If commercially printed grids are used in graph drawings, author should be sure printer's ink is of a color that will reproduce. All half-tone illustrations (photographs, wash, airbrush, or pencil renderings, etc.) should be clean and ready to reproduce. Photographs should be glossy prints. Call-outs or labels should be marked on a registered tissue overlay, not on the illustration itself. No illustration should be larger than 8 x 10 inches.

Copies can be purchased from
THE INSTITUTE OF RADIO ENGINEERS
1 East 79 St., New York 21, N.Y.

PRICE PER COPY: members of the Professional Group on Antennas and Propagation, $2.20; members of the IRE, $3.30; nonmembers, $6.60.

ANNUAL SUBSCRIPTION PRICE: IRE members, $8.50; Colleges and public libraries, $10.00; nonmembers, $17.00.

ed as second-class matter, at the post office at Menasha, Wisconsin, under the act of August 24, 1912. ance for mailing at a special rate of postage is provided for in the act of February 28, 1925, embodied in Paragraph 4, Section 412, P. L. & R., authorized October 26, 1927.

news and views

PGAP Survey

The membership was surveyed by a post card questionnaire to obtain better information on needs and interests. The number of returns was gratifyingly high, but, unfortunately, a good many of the group did not complete the questionnaire in the desired form so that they could not be tallied. A sufficient number of members did respond correctly and the tally of the correctly filled-in questionnaires probably gives a statistically satisfactory result.

The post card questionnaire is reproduced here for convenience in checking the tabulated returns:

PGAP Survey

(Number Replies in Order of Importance)

Field of Interest:
Antennas
Propagation
Microwave Tech.
Microwave Optics
Radio Astronomy

Prof. Group Membership:
Antennas & Propagation
Microwave Theory & Tech.
List Other Prof. Groups
Comm. Systems
Telemetry & Remote Control

Paper Pref. for Transactions:
Experimental Type
Analytical Type
Applications Type

Meeting Preferences:
More-Less National Meetings
More-Less Local Chapter Mtgs.

Local Chapter Membership:
Do you support your Local PGAP Chapter? Name Local.
Do you support any Local IRE Group?

The results of the survey are shown in Table I. Under Field of Interest, a tally for first through fifth place is shown for each field. Thus under Radio Astronomy, five people indicated this field as their first preference, thirty-one people indicated this to be their second preference, and so forth. Ideally, if everyone filled in all numbers from one to five on the questionnaire, the vertical and horizontal totals should be the same. This is not so because of the interesting display of individuality in the replies, where some filled in perhaps two out of the five boxes, etc.

However, it does appear that Antennas, Propagation, and Microwave Techniques, in order, represent the greatest interest to our group with an interesting de-emphasis on Radio Astronomy.

TABLE I
PGAP Survey

	1	2	3	4	5
Field of Interest					
Antennas	249	161	42	8	2
Propagation	157	130	98	68	22
Microwave Tech.	84	98	90	44	24
Microwave Optics	18	50	104	90	56
Radio Astronomy	5	31	60	67	130
Paper Preference					
Experimental	113	190	59		
Analytical	178	90	164		
Applications	155	153	148		
Meeting Preference					
More National Meetings	157				
Less National Meetings	94				
More Local Chapter Meetings	189				
Less Local Chapter Meetings	42				

Under Papers Preference, there is no pronounced trend. While analytical papers appear to lead as first preference, about as many people also put such papers in third place. One can at least say that people are not neutral about analytical type papers; they are either for them or against them.

Finally, there is a fairly definite preference for more national meetings and more local chapter meetings.

In a general way, the results are in line with the thinking of the Administrative Committee as to their preconceived feelings of the interests and needs of the membership. The Administrative Committee is thus encouraged to proceed in its previously set general directions, particularly insofar as more national and local chapter meetings are concerned.

Local Chapters

An application for a new section of PGAP at San Diego has been approved by the Administrative Committee. On approval of the IRE Executive Committee this group can officially get under way.

Formation of a PGAP Chapter in Columbus, Ohio is under consideration. Dr. D. D. Knig of Johns Hopkins

University is the member of the National Administrative Committee responsible for the liaison and promotion of local chapters and he will be glad to assist in any way possible.

Washington Chapter—The Washington Chapter looks as though it will become quite active. A regular meeting was held on March 6, 1956 at George Washington University with speaker Dr. D. D. King of Johns Hopkins who presented a paper on recent work at his laboratory on millimeter wave research. The agenda for the meeting appears quite interesting. Unfortunately, a report of the discussion is not available at this writing. The agenda, however, is reproduced below:

Business

1) A nominating committee will be appointed to consider candidates for chapter offices next year.

2) A symposium is being planned for this Fall and would possibly be organized along the same lines as the Scatter Symposium which was held last November. Some of the groundwork for this has been started and with the cooperation of chapter membership there is no reason why this area cannot be the site for another outstanding symposium.

3) There are a number of items relating to the overall policies of PGAP, particularly with regard to meetings and publications, which should be of interest to all members. This might well be the source of a lively and beneficial discussion.

A Reminder

Our meeting this Tuesday night is next to the last which will be held this year. There is much to be done toward outfitting the chapter for its intended purpose, which is, primarily, to serve the needs of the membership in advancing our technical knowledge in the field of antennas and propagation.

Problems always exist relative to the procurement of interesting speakers, balancing of the different interests among members, timing of the meetings, and cooperation with other PG chapters. I most sincerely urge that you make it a definite point to attend this next meeting and express your views so that sound policies can be formulated for direction of our efforts next year.

—Coleman Goatley, *Chairman*

Meetings

WESCON, August 21–24, 1956

The 1956 WESCON will be held in Los Angeles, California this year and as in the past the PGAP will sponsor the technical sessions on antennas and propagation.

Paper summaries are required by April 15, 1956 and should be addressed to:

Dr. R. S. Elliott
Hughes Research Laboratories
Culver City, California

Authors will be notified of the disposition of their papers by May 15.

Since a Convention Record will be published along lines parallel to the IRE Convention Record, it will be necessary to have reproducible copies of all papers accepted in Dr. Elliott's hands by June 15, 1956.

The recommended format for the reproducible copy will be furnished at the same time the disposition of the papers is announced.

WESCON Schedule for Papers

April 15—Summaries to Dr. R. S. Elliott.
May 15—Author notification.
June 15—Reproducible copies to Dr. Elliott.

Correspondence

The following is an excerpt from a letter from John Smyth to George Sinclair containing some thoughts worthy of discussion:

"At no time has a reviewer been identified to an author. When the paper is acceptable, with or without some revision, the reviewer's remarks are sent along to the author as a service to him. When the paper appears in print the reviewer as well as anyone else can send along detailed comments for publication. The reviewer's comments are the basis for accepting or rejecting papers for publication; they are also an aid to the author in that they frequently afford him an opportunity to improve on his paper, correcting errors, filling in desired details and the like. Rumsey's statement, 'This material may be appended to the published paper (as comments by the referee) if the author so desires' is quite acceptable. Or, in the event the author wishes to ignore some of the referee's comments, the reviewer can always write a letter to the editor about the paper once it is in print.

I think it would improve our Transactions if more people sent along their remarks about published papers for the Communications section. One might go so far as to say the English system of simultaneously publishing the comments of the referee and rebuttal by the author is better than the way we handle our reviews. In fact, I personally like it.

Rumsey's idea of handling reviews is certainly worthy of considerable thought and discussion and I recommend that it be placed high on the agenda of the next Administrative Committee meeting."

Calendar of Events

Apr. 30–May 3, 1956: URSI Spring Meeting, National Bureau of Standards, Washington, D. C.

May 1–3, 1956: IRE-RETMA-AIEE-WCEMA Electronic Components Symposium, U. S. Department of Interior, Washington, D. C.

May 14–16, 1956: National Aeronautical and Navigational Conference, Hotel Biltmore, Dayton, Ohio.

May 21–22, 1956: Symposium on Reliable Applications of Electron Tubes, University of Pennsylvania, Philadelphia, Pa.

June 4–6, 1956: Second Annual Radome Symposium, Ohio State University, Columbus, Ohio.

Aug. 20–21, 1956: National Telemetering Conference, Biltmore Hotel, Los Angeles, Calif.

Aug. 21–24, 1956: IRE-West Coast Electronic Manufacturers' Association, WESCON, Biltmore Hotel, Los Angeles, Calif.

Sept. 11–12, 1956: Second RETMA Conference on Reliable Electrical Connections, University of Pennsylvania, Philadelphia, Pa.

Sept 17–21, 1956: Instrument-Automation Conference and Exhibit, Coliseum, New York, N. Y.

Sept. 24–25, 1956: Industrial Electronics Symposium, Manger Hotel, Cleveland, Ohio.

Oct. 1–3, 1956: National Electronics Conference, Chicago, Ill.

Oct. 1–3, 1956: Canadian IRE Convention and Exposition, Automotive Building, Exhibition Park, Toronto, Canada.

Oct. 8–9, 1956: Second Annual Symposium on Aeronautical Communications, Utica, N. Y.

Oct. 10–12, 1956: Symposium on Applications of Optical Principles to Microwaves, Washington, D. C.

contributions

Variational Principles for Electromagnetic Resonators and Waveguides*

A. D. BERK†

Summary—**Variational expressions are presented for the propagation constants of a waveguide and for the resonant frequencies of a cavity directly in terms of the field vectors for situations not covered in the literature. These situations occur when the electromagnetic problem is not susceptible of a scalar formulation and are typified by the presence of inhomogeneous or anisotropic matter. The variational expressions are applied to several illustrative examples and the results are compared to known exact solutions. The variational expressions are also applied to the derivation of certain perturbation formulas, some of which were previously obtained by different methods.**

Introduction

KNOWN variational expressions for the resonance frequencies of a resonator are restricted to the special case when the electromagnetic field is derivable from a single scalar function which satisfies the Helmholtz equation.[1] When the substance within a cavity is inhomogeneous or anisotropic, such a scalar formulation is inadequate. The need for vector variational principles is thus apparent. It is the purpose of the first part of this paper to present such variational formulas for resonant frequencies directly in terms of the field vectors. In problems of propagation through anisotropic or inhomogeneous media, variational expressions for the propagation constant are also of interest. The second part of this paper is concerned with the derivation of such expressions. The variational expressions for resonant frequencies and those for propagation constants appearing in this paper are to be distinguished from variational expressions for reflection coefficients, scattering amplitudes or impedance matrices derived by other authors.[1–3] In the third part, the variational formulas are applied to illustrative specific examples and to the derivation of some perturbation expressions.

Variational Principles for Resonant Frequencies

E-Field Formulation

Consider a resonator with perfectly conducting walls enclosing a medium of permittivity ϵ and permeability μ. Both μ and ϵ may be tensors[4] and functions of position. Let a resonant angular frequency be ω and let the corresponding electromagnetic field be characterized by the vectors E and H. The following is then asserted to be a variational expression for ω, provided ϵ and μ are hermitian; *i.e.*, provided no losses are present:

$$\omega^2 = \frac{\int (\text{curl } E^*) \cdot \mu^{-1} \cdot (\text{curl } E) dv}{\int E^* \cdot \epsilon \cdot E dv}. \tag{1}$$

* Manuscript received by the PGAP, August 11, 1955; revised manuscript received January 27, 1956. The research reported in this paper is based on a chapter of a doctoral thesis submitted by the author to the Massachusetts Institute of Technology in September, 1954 and was supported in part by the Signal Corps, the Office of Scientific Research, Air Research and Development Command, and the Office of Naval Research.

† Hughes Aircraft Co., Culver City, Calif.

[1] P. M. Morse, and H. Feshbach, "Methods of Theoretical Physics," McGraw-Hill Book Co., New York, N. Y., p. 1112; 1953.

[2] D. S. Saxon, Notes on Lectures by J. Schwinger, "Discontinuities in Waveguides," Radiation Laboratory, M.I.T.; February, 1945.

[3] H. Levine and J. Schwinger, "On the theory of diffraction by an aperture in an infinite screen plane," *Phys. Rev.*, vol. 74, pp. 958–974; October, 1948.

[4] For simplicity, no special symbols will be used to indicate the tensor character of these quantities.

The integrals are over the volume of the resonator, μ^{-1} is the inverse of μ, E^* is the complex conjugate of E. To prove this assertion, we must show that those field configurations E and E^* which render ω^2 stationary are solutions of

$$\text{curl}\,(\mu^{-1}\cdot\text{curl}\,E) - \omega^2\epsilon\cdot E = 0^{5} \tag{2}$$

and of the complex conjugate of (2) and have vanishing tangential components at the boundary. This is indeed the case, for, on varying E and E^* in (1) we obtain, after utilizing the hermitian character of ϵ and μ, the following expression for the variation of ω^2:

$$\left(\int E^*\cdot\epsilon\cdot E dv\right)\delta\omega^2 = \int \delta E^*\cdot[\text{curl}\,(\mu^{-1}\cdot\text{curl}\,E) - \omega^2\epsilon\cdot E]dv - \oint \delta E^*\cdot(n\times\mu^{-1}\cdot\text{curl}\,E)ds + \int \delta E\cdot[\text{curl}\,(\mu^{-1*}\cdot\text{curl}\,E^*) - \omega^2\epsilon^*\cdot E^*]dv - \oint \delta E\cdot(n\times\mu^{-1*}\cdot\text{curl}\,E^*)ds. \tag{3}$$

The second and fourth integrals are over the boundary of the cavity. Their appearance is due to the utilization of the vector identity

$$\int A\cdot\text{curl}\,B dv = \int B\cdot\text{curl}\,A dv + \oint n\cdot(B\times A)ds, \tag{4}$$

where n is the outward normal unit vector. The variation of ω^2 will vanish, provided E satisfies (2), E^* satisfies the complex conjugate of (2) and provided the surface integrals in (3) vanish. The latter condition can be satisfied only if $n\times\delta E$ and its complex conjugate vanish over the boundary, since $n\times(\mu^{-1}\cdot\text{curl}\,E)$, being proportional to the tangential component of the magnetic field, cannot vanish over the complete boundary. Eq. (1) is thus a variational formulation of the problem defined by (2) and the boundary condition $n\times E=0$. Admissible trial fields must have vanishing tangential components at the boundary, must be continuous together with their first derivatives and must possess finite second derivatives everywhere in the cavity except at surfaces where ϵ and μ are discontinuous. At such surfaces $n\times E$ and $n\times(\mu^{-1}\cdot\text{curl}\,E)$ must be continuous. Eq. (1) can be modified so that trial vectors E are *not* required to satisfy the boundary condition $n\times E=0$ at the wall of the resonator. This can be achieved following a known general method[6] by adding appropriate terms to the numerator of (1) as follows:

$$\omega^2 = \frac{\int(\text{curl}\,E^*)\cdot\mu^{-1}\cdot(\text{curl}\,E)dv - \oint n\cdot[E\times(\mu^{-1*}\cdot\text{curl}\,E^*)]ds - \oint n\cdot[E^*\times(\mu^{-1}\cdot\text{curl}\,E)]ds}{\int E^*\cdot\epsilon\cdot E dv}, \tag{5}$$

or, on combining the first and third terms in the numerator,

$$\omega^2 = \frac{\int E^*\cdot\text{curl}\,(\mu^{-1}\cdot\text{curl}\,E)dv - \oint n\cdot[E\times(\mu^{-1*}\cdot\text{curl}\,E^*)]ds}{\int E^*\cdot\epsilon\cdot E dv}. \tag{6}$$

When the distribution of matter within the cavity is discontinuous, (1) can be further modified so that trial fields will *not* be required to have continuous tangential components of E and of $(\mu^{-1}\cdot\text{curl}\,E)$. The modification consists of adding to the numerator of (1) the term

$$-\int n\cdot[E_+^*\times(\mu_+^{-1}\cdot\text{curl}\,E_+) - E_-^*\times(\mu^{-1}\cdot\text{curl}\,E_-)]ds$$

$-$ complex conjugate,

where the subscripts $+$, $-$ refer to values on opposite sides of the surface of discontinuity and the integrals are over such a surface.

The passage from (1) to (6), enables one to expand the class of admissible trial functions. The value of such a procedure will be demonstrated in the section on Applications, where trial functions outside of the class associated with (1) will be utilized.

Finally, (1) can be modified to apply when the boundary condition at the walls is

$$n\times(\mu^{-1}\cdot\text{curl}\,E) = -j\omega\Upsilon\cdot E_t \tag{7}$$

where $(j\omega\Upsilon)$ is a hermitian two dimensional dyadic and E_t is the tangential component of E. The term to be added to the numerator of (1) is, in this case, $\oint E_t^*\cdot(j\omega\Upsilon)\cdot E_t ds$. A practical example where (7) applies, is at the junction of a cavity and a tuning stub.

[5] This equation is the result of eliminating the magnetic field between Maxwell's equations.

[6] R. Courant and D. Hilbert, "Methods of Mathematical Physics," First English Translation, Interscience Publishers, pp. 208–211, 1953. See also reference 1, pp. 1131–33.

H-Field Formulation

A variational expression similar to (1) but in terms of the magnetic field can be obtained by interchanging ϵ and μ and by replacing E by H:

$$\omega^2 = \frac{\int (\text{curl } H^*) \cdot \epsilon^{-1} \cdot (\text{curl } H) dv}{\int H^* \cdot \mu \cdot H dv}. \tag{8}$$

The distinguishing feature of (8) is that trial vectors H are *not* required to satisfy the proper boundary condition, $n \times (\epsilon^{-1} \cdot \text{curl } H) = 0$, at the wall. However, the differentiability and continuity conditions are the same as those for E in connection with (1).

Eqs. (1) and (8) reduce properly to the variational principle of reference 1 in the special case when the electromagnetic problem can be expressed in terms of a single scalar field.

Mixed-Field Formulation

In contrast to the preceding formulas which are in terms of either the electric vector or the magnetic vector, the following variational expression is in terms of both field vectors.

$$\omega = j \frac{\int H^* \cdot \text{curl } E dv = \int E^* \cdot \text{curl } H dv}{\int E^* \cdot \epsilon \cdot E dv + \int H^* \cdot \mu \cdot H dv}, \tag{9}$$

where ϵ, μ are again assumed to be hermitian. That (9) is indeed a variational expression can readily be verified by computing the variation of ω and observing that the latter vanishes, provided E, H satisfy Maxwell's equations, provided E^*, H^* satisfy the complex conjugate of Maxwell's equations, and provided the trial electric vectors have vanishing tangential components at the boundary. Thus, admissible trial E-vectors must be continuous, must possess first derivatives and their tangential components at the boundary must vanish. Admissible trial H-vectors are subject to the same continuity and differentiability conditions as the E-vectors, but they are not required to satisfy any particular boundary condition. If matter is discontinuously distributed within the cavity, then trial vectors E and H must have continuous tangential components at the surfaces of discontinuity. The latter restriction and the restriction of vanishing trial tangential E at the walls of the cavity can be eliminated by the addition of appropriate terms to the numerator of (9). For example, in the following variational expression

$$\omega = j \frac{\int H^* \cdot \text{curl } E dv - \int E^* \cdot \text{curl } H dv - \oint n \cdot (E \times H^*) ds}{\int E^* \cdot \epsilon \cdot E dv + \int H^* \cdot \mu \cdot H dv} \tag{10}$$

both trial vectors E and H are unrestricted at the wall.

Variational Formulas for the Propagation Constant

Mixed-Field Formulation

Consider a waveguide with perfectly conducting walls, possibly enclosing anisotropic matter whose distribuion, may be a function of the transverse coordinates but not of the coordinate along the direction of propagation. If z is this coordinate, the field vectors may be expressed as $\mathcal{E}(x, y)e^{-j\gamma z}$, $\mathcal{H}(x, y)e^{-j\gamma z}$, where γ is the propagation constant. $\mathcal{E}$ and $\mathcal{H}$ are three dimensional vectors depending on x and y only and they satisfy the following relations obtained by substituting the field vectors in Maxwell's equations:

$$\text{curl } \mathcal{E} + j\omega\mu \cdot \mathcal{H} = \gamma j a_z \times \mathcal{E}, \tag{11}$$

$$\text{curl } \mathcal{H} - j\omega\epsilon \cdot \mathcal{E} = \gamma j a_z \times \mathcal{H}, \tag{12}$$

where a_z is the unit vector in the z-direction.

Premultiplying (11) by $\mathcal{H}^*$, (12) by $\mathcal{E}^*$, integrating over the cross section of the waveguide and subtracting we obtain the following variational expression:

$$\gamma = \frac{\omega \int \mathcal{E}^* \cdot \epsilon \cdot \mathcal{E} ds + \omega \int \mathcal{H}^* \cdot \mu \cdot \mathcal{H} ds + j \int \mathcal{E}^* \cdot \text{curl } \mathcal{H} ds - j \int \mathcal{H}^* \cdot \text{curl } \mathcal{E} ds}{\int \mathcal{H} \cdot {}^* a_z \times \mathcal{E} ds - \int \mathcal{E}^* \cdot a_z \times \mathcal{H} ds}. \tag{13}$$

That (13) is indeed a variational expression can be shown by evaluating the variation of γ and observing that the latter vanishes provided $\mathcal{E}$, $\mathcal{H}$ satisfy (11) and (12) and provided the tangential component of $\mathcal{E}$ vanishes at the walls of the waveguide. Thus, trial fields $\mathcal{E}$, $\mathcal{H}$ must be continuous, and must possess first derivatives throughout the cavity. At the boundary, the tangential component of $\mathcal{E}$ must vanish, but $\mathcal{H}$ is arbitrary. When discontinuities are present in the distribution of matter within the cavity, the tangential components of both $\mathcal{E}$ and $\mathcal{H}$ must be continuous at the surfaces of discontinuity.

In the following modified form of (13) both $\mathcal{E}$ and $\mathcal{H}$ are arbitrary at the boundary.

$$\gamma = \frac{\omega \int \mathcal{E} \cdot {}^{*}\epsilon \cdot \mathcal{E} ds + \omega \int \mathcal{H}^{*} \cdot \mu \cdot \mathcal{H} ds + j \int \mathcal{E}^{*} \cdot \text{curl } \mathcal{H} ds - j \int \mathcal{H}^{*} \cdot \text{curl } \mathcal{E} ds - j \oint n \cdot (\mathcal{E} \times \mathcal{H}^{*}) dl}{\int \mathcal{H}^{*} \cdot a_z \times \mathcal{E} ds - \int \mathcal{E}^{*} \cdot a_z \times \mathcal{H} ds} . \quad (14)$$

The last integral in the numerator of (14) is over the boundary of the cross section of the waveguide.

$\mathcal{E}$-Field Formulation

If $\mathcal{H}$ is eliminated between (11), (12), the result premultiplied by $\mathcal{E}^{*}\cdot$ and integrated over the cross section, the following is obtained:

$$\gamma^2 \int (a_z \times \mathcal{E}^{*}) \cdot \mu^{-1} \cdot (a_z \times \mathcal{E}) ds$$

$$-j\gamma \int [(\text{curl } \mathcal{E}^{*}) \cdot \mu^{-1} \cdot (a_z \times \mathcal{E}) - (a_z \times \mathcal{E}^{*}) \cdot \mu^{-1} \cdot (\text{curl } \mathcal{E})] ds$$

$$+ \int (\text{curl } \mathcal{E}^{*}) \cdot \mu^{-1} \cdot (\text{curl } \mathcal{E}) ds - \omega^2 \int \mathcal{E}^{*} \cdot \epsilon \cdot \mathcal{E} ds = 0. \quad (15)$$

Varying $\mathcal{E}$, $\mathcal{E}^{*}$ and γ in (15) it can be shown that γ remains stationary provided $\mathcal{E}$, $\mathcal{E}^{*}$ satisfy

$$\gamma^2 a_z \times (\mu^{-1} \cdot a_z \times \mathcal{E}) + j\gamma [\text{curl } (\mu^{-1} \cdot a_z \times \mathcal{E}) + a_z \times (\mu^{-1} \cdot \text{curl } \mathcal{E})] - \text{curl } (\mu^{-1} \cdot \text{curl } \mathcal{E}) + \omega^2 \epsilon \cdot \mathcal{E} = 0 \quad (16)$$

and the conjugate of (16). Eq. (16) is the result of eliminating $\mathcal{H}$ between (11) and (12). Hence, it is the proper "Euler" equation. Solving (15) for γ we obtain a variational expression for the propagation constant.

Admissible trial vectors $\mathcal{E}$ must have vanishing tangential components at the walls, must be continuous together with their first derivatives and must possess second derivatives except at surfaces of discontinuity of the medium enclosed within the waveguide. At such surfaces, $\mathcal{E}$ and $(\mu^{-1} \cdot \text{curl } \mathcal{E} - j\gamma\mu^{-1} \cdot a_z \times \mathcal{E})$ must have continuous tangential components.

$\mathcal{H}$-Field Formulation

If in (15) μ and ϵ are interchanged and $\mathcal{E}$ is replaced by $\mathcal{H}$, a new variational equation results. Admissible trial vectors $\mathcal{H}$ must satisfy the same continuity and differentiability conditions as does $\mathcal{E}$ in connection with (15), but they are arbitrary at the boundary of the waveguide.

The variational principles referred to in the preceding sections can be modified so that trial functions will be less sensitive to boundary conditions. The method is similar to that employed in deriving (5) and (10) and will not be repeated here.

Applications

We have, thus far, offered several variational expressions for both the resonant frequencies of a cavity and the propagation constant of a waveguide. We shall now illustrate them by a number of examples. For a particular problem, the choice of a certain variational formula in preference to others largely depends on physical considerations. For example, if the configuration of the electric field can be "guessed" more readily than that of the magnetic field, then it is sensible to use a formula in terms of the electric vector only.

The successful choice of trial fields also depends on our familiarity with the physical aspects of the problem at hand. It is important to remember that the conditions of admissibility of trial functions are minimum requirements and that one should make every effort to select trial fields which satisfy as many of the known features of the solution as possible. In particular, one should attempt to devise trial fields which, at surfaces of discontinuity, not only have continuous tangential components, but continuous normal components of $(\epsilon \cdot E)$ and $(\mu \cdot H)$ as well. This is because the last set of boundary conditions does *not* follow from the first set unless the trial fields satisfy Maxwell's equations.

The first three of the following examples are intended to illustrate the advantages of approximate calculations based on the variational expressions of this paper over exact treatments. In the remaining examples, the variational expressions are applied to the derivation of certain perturbation formulas, some of which have already appeared in the literature by different authors.

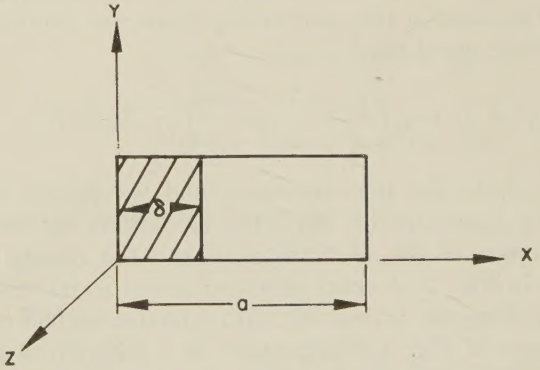


Fig. 1—Cross section of a waveguide with a dielectric slab.

Specific Examples

Let it be required to find the cut-off frequency ω_c of the fundamental mode of a rectangular waveguide of width a in which a dielectric slab of thickness δ is placed adjacent to one of the walls as shown in Fig. 1. The slab has a permittivity ϵ different from ϵ_0, the permittivity of air. Its permeability, however, is assumed to be μ_0,

the permeability of air. Now, the cut-off frequency of a waveguide is the resonant frequency of a two dimensional cavity defined by the cross section of the guide. Let us then use (1) with $E = a_y \sin(\pi x/a)$ as a trial field. The result is

$$\omega_c^2 \mu_0 \epsilon_0 = \left(\frac{\pi}{a}\right)^2 \cdot \left[1 + \frac{(\epsilon - \epsilon_0)}{\epsilon_0}\left(\frac{\delta}{a} - \frac{1}{2\pi}\sin\frac{2\pi\delta}{a}\right)\right]^{-1}. \quad (17)$$

Upon introducing the cut-off wavelength λ_c related to ω_c by $\omega_c^2 \mu_0 \epsilon_0 = (2\pi/\lambda_c)^2$ and computing a/λ_c for various thicknesses of the dielectric slab and for a dielectric constant of 2.45 (polystyrene) Table I is obtained. The results are compared to the exact values given in a separate column and read from published curves.[7] Note the simplicity of (17) and its possible use as a design equation. The exact treatment requires the solution of a transcendental equation.

TABLE I

CUT-OFF WAVELENGTH OF A RECTANGULAR WAVEGUIDE WITH A DIELECTRIC SLAB

δ/a	a/λ_c (exact)	a/λ_c (approximate)
0.000	0.500	0.500
0.167	0.485	0.486
0.286	0.450	0.455
0.500	0.375	0.383
0.600	0.355	0.352
1.00	0.318	0.320

Next we shall calculate the propagation constant of the fundamental mode of a rectangular waveguide with a dielectric slab located symmetrically as shown in Fig. 2. The slab has a dielectric susceptibility $\chi_e = (\epsilon/\epsilon_0) - 1$ and a magnetic susceptibility zero. Utilizing the variational expression (16) and taking $E = a_y \sin(\pi x/a)$, as a trial field, we obtain

$$(\lambda/\lambda_g)^2 = 1 + \chi_e\left(\frac{\delta}{a} + \frac{1}{\pi}\sin\frac{\pi\delta}{a}\right) - .25(a/\lambda)^{-2}, \quad (18)$$

where λ, λ_g are the free-space and waveguide wavelengths, respectively. Eq. (18) is plotted for various thicknesses of the dielectric slab and the results compared in Fig. 2 to exact solutions given in reference 5.

A rectangular waveguide with a ferrite slab off center as shown in Fig. 3 is important as a differential phase shifter. The differential propagation constant is defined as the difference of the propagation constants in the two opposite directions, $\gamma_+ - \gamma_-$. This difference is a consequence of the tensor susceptibility of the ferrite, which for the configuration shown in Fig. 3, is of the form:

$$\chi_m = \begin{bmatrix} x & 0 & -j\kappa \\ 0 & 0 & 0 \\ j\kappa & 0 & x \end{bmatrix}.$$

[7] C. G. Montgomery, R. H. Dicke, and E. M. Purcell, "Principles of Microwave Circuits," M.I.T. Rad. Lab. Series, vol. 8, p. 387; 1948.

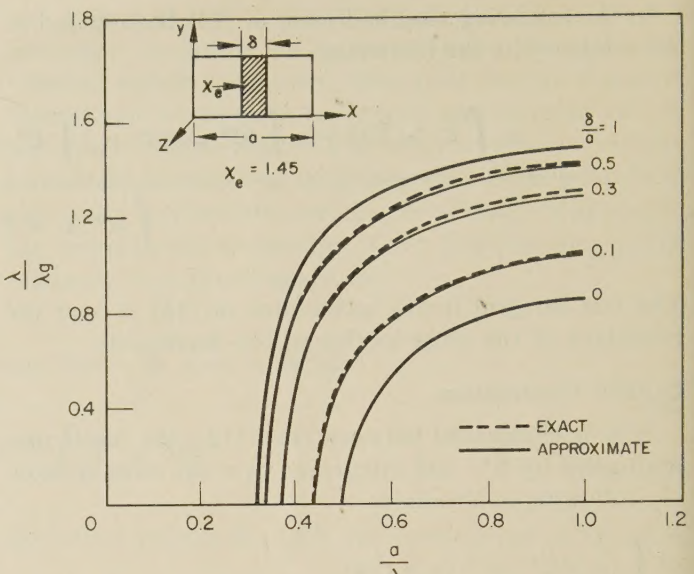


Fig. 2—Variation of propagation constant vs frequency for slabs of various widths in a rectangular waveguide.

An important quantity is the displacement of the slab from the wall, d, for which the differential propagation constant is a maximum. To determine this quantity, we must first determine the propagation constant. Let us utilize, for this purpose, (13) with a trial field selected as follows: For the $\mathcal{E}$-field we take,

$$\mathcal{E} = a_y\left(\sin\frac{\pi x}{a} + A\sin\frac{2\pi x}{a}\right), \quad (19)$$

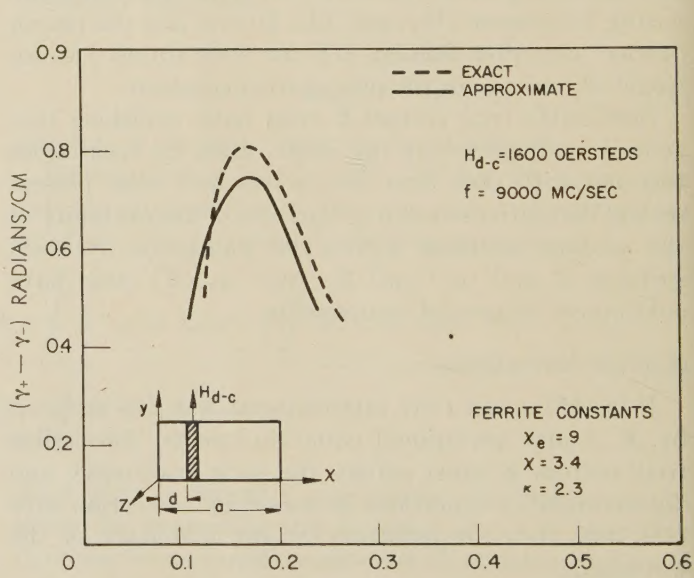


Fig. 3—Differential propagation constant as a function of the location of a ferrite slab in a rectangular waveguide.

where A is a variational parameter. We determine $\mathcal{H}$ by substituting (19) into (11) where for this equation only μ is assumed to be a scalar and equal to μ_0 and γ is assumed to equal the propagation constant of the fundamental mode with the ferrite slab replaced by a dielec-

tric one[8] having the same permittivity as that of the ferrite, but with a permeability equal to that of the free space. Substituting $\mathcal{E}$ and $\mathcal{H}$ in (13), determining A from the condition $\partial\gamma/\partial A = 0$, and substituting the value of A, thus determined back in (13), γ is obtained. Note that the immediately preceding procedure involves several simplifying assumptions and it amounts essentially to utilizing the variational principle (13) subject to the constraint (11). Details have been omitted because they are lengthy. The differential propagation constant plotted as a function of the distance d and for specific values of slab thickness and tensor permeability is shown in Fig. 3 and compared with the exact solution obtained by Lax, Button, and Roth.[9] Note the agreement in the location of the maximum. The ratio of slab thickness to the width of the guide is 0.044. For ratios appreciably less than 1 per cent, we can simply substitute for $\mathcal{E}$ and $\mathcal{H}$ in (13) the expressions valid for an empty waveguide taking into account, however, the continuity of the normal component of the magnetic flux density at the surface of the slab. The result is

$$\gamma_+ - \gamma_- = 2\frac{\pi}{a}\frac{\delta}{a}\frac{\kappa}{1+x}\sin\frac{2\pi d}{a}, \tag{20}$$

where δ is the thickness of the slab.

This is the expression Lax, Button, and Roth obtained from an expansion of a transcendental equation.[9]

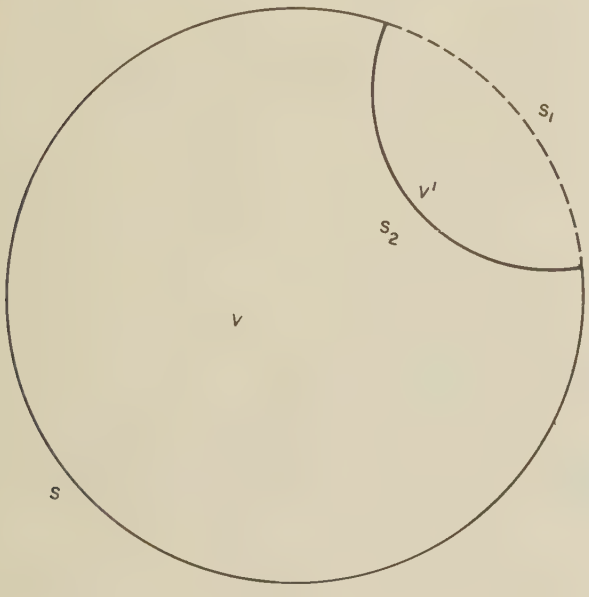


Fig. 4—Cavity with deformed boundary.

Perturbation Formulas

Cavity Surface Perturbation: For a general application of (6) consider the following problem: A resonator bounded by conducting surfaces S, S_1 is deformed so that surface S_1 is replaced by S_2 as shown in Fig. 4. If the field configuration and the resonant frequency for the undeformed cavity are characterized by E, H, ω an approximate expression for the frequency of the deformed cavity is obtained by substituting E as the trial field in (6). Note that E violates the boundary condition at S_2 but it is an admissible trial field because of the presence of the surface integral in (6). The result is

$$\omega_1^2 = \omega^2\left[1 + \frac{\int_{v'}[H^*\cdot\mu\cdot H - E^*\cdot\epsilon\cdot E]dv}{\int_v E^*\cdot\epsilon\cdot E dv}\right], \tag{21}$$

where v' is the volume bounded by S_1 and S_2. When μ and ϵ are scalars (21) becomes identical to a formula derived by Slater by a different method.[10]

Waveguide Wall Perturbation: Let $\mathcal{E}$, $\mathcal{H}$, γ represent the known solution for a certain waveguide. Suppose, now, that the cross section is deformed so that s is the cross-sectional surface between the original and final cross sections and let γ_1 be the new propagation constant. Substituting $\mathcal{E}$, $\mathcal{H}$ as trial vectors in (14) we obtain

$$\gamma_1 - \gamma = \omega\frac{\int_s \mathcal{H}^*\cdot\mu\cdot\mathcal{H}ds - \int_s \mathcal{E}^*\cdot\epsilon\cdot\mathcal{E}ds}{\int_S \mathcal{H}^*\cdot a_z\times\mathcal{E}ds - \int_S \mathcal{E}^*\cdot a_z\times\mathcal{H}ds}. \tag{22}$$

The integrals of the denominator are over the cross-sectional surface of the deformed waveguide.

Cavity Volume Perturbation: Let a resonator be characterized by ω, E, H, when the permittivity and the permeability of the medium are ϵ, μ. Let ω_1 be the resonant frequency when ϵ, μ change to ϵ_1, μ_1, the boundary of the resonator remaining the same. Substituting E, H as trial fields in (9) we get

$$\omega_1 = \omega\frac{\int H^*\cdot\mu\cdot H dv + \int E^*\cdot\epsilon\cdot E dv}{\int H^*\cdot\mu_1\cdot H dv + \int E^*\cdot\epsilon_1\cdot E dv} \tag{23}$$

or, forming the relative frequency shift,

$$\frac{\Delta\omega}{\omega_1} = \frac{\omega_1 - \omega}{\omega_1} = -\frac{\int H^*\cdot(\mu_1-\mu)\cdot H dv + \int E^*\cdot(\epsilon_1-\epsilon)\cdot E dv}{\int H^*\cdot\mu\cdot H dv + \int E^*\cdot\epsilon\cdot E dv}. \tag{24}$$

Eq. (24) can also be obtained from a well known formula.[11]

[8] A rectangular waveguide with a dielectric slab can be approximately determined by the methods discussed in this paper.

[9] B. Lax, K. J. Button, and L. M. Roth, "Ferrite phase shifters in rectangular waveguide," *J. Appl. Phys.*, vol. 25, pp. 1413–1421; November, 1954.

[10] J. C. Slater, "Microwave Electronics," D. Van Nostrand Company, New York, N. Y., p. 81, 1950.

[11] H. B. G. Casimir, Philips Res. Rep., vol. 6, pp. 162–182, 1951. Essentially the same formula was also derived by Bethe and Schwinger in a Rad. Lab. Report of M.I.T., March 4, 1943.

Waveguide Volume Perturbation: Let a known solution be γ, $\mathcal{E}$, $\mathcal{H}$ for a certain waveguide when the permittivity and permeability of the medium are ϵ, μ. Let us find γ_1, the propagation constant when ϵ, μ are changed to ϵ_1, μ_1, the boundary of the waveguide remaining the same. Substituting $\mathcal{E}$, $\mathcal{H}$ as trial fields in (13) we obtain

$$\gamma_1-\gamma=\omega\frac{\int\mathcal{E}^*\cdot(\epsilon_1-\epsilon)\cdot\mathcal{E}ds+\int\mathcal{H}^*\cdot(\mu_1-\mu)\cdot\mathcal{H}ds}{\int\mathcal{H}^*\cdot a_z\times\mathcal{E}ds-\int E^*\cdot a_z\times\mathcal{H}ds}. \quad (25)$$

Eq. (25) can also be derived from a formula first derived, but not published, by B. Lax of Lincoln Laboratory, M.I.T.[12]

Acknowledgment

The author is indebted to R. B. Adler, L. J. Chu, both of Massachusetts Institute of Technology, to Dr. B. Lengyel of the Hughes Research Laboratories, and to Dr. B. Lax of Lincoln Laboratory, M.I.T., for important criticism and discussions.

Discussion

Comments by the Referee

It should be noted that formulas similar to those given by Berk, but applicable to the general lossy case, can be derived by using the reaction concept.[13] The method can be illustrated by considering the problem of calculating the resonant frequency (and bandwidth, if the loss is not zero) in terms of an assumed distribution of E in a closed cavity. To apply the reaction concept we have to find some source which would generate the assumed distribution. Let the source distribution be taken as a volume distribution of electric dipoles J plus a surface distribution of magnetic dipoles K_s where

$$\nabla\times H=(j\omega\epsilon+\sigma)E+J \quad (26)$$

$$\nabla\times E=-j\omega\mu H \quad (27)$$

$$\therefore\ J=-(j\omega\epsilon+\sigma)E+\frac{j}{\omega}\nabla\times(\mu^{-1}\nabla\times E) \quad (28)$$

and

$$K_s=n\times E \text{ on } S \quad (29)$$

where S is the boundary of the cavity and n a unit normal vector pointing outwards. Inspection of the above equations shows that the combination of J and K_s generates the assumed E in the cavity, for Maxwell's equations are satisfied at every interior point and the boundary conditions on S are satisfied.

Using the method and notation of the reaction concept[13] this leads to the equation

$$\langle a, a\rangle=0 \quad (30)$$

for the resonant frequency (and bandwidth) where "a" represents the approximate source consisting of the combination of J and K_s. [It can be verified that the value of ω obtained from (30) is stationary for variations of the assumed E about the correct distribution (cf. ref. 1).] Written out in detail (30) gives:

$$\begin{aligned}0&=\iiint_V J\cdot Edv-\iint_S K_s\cdot HdS\\&=-\iiint_V E\cdot(j\omega\epsilon+\sigma)Edv\\&\quad+\iiint_V\frac{j}{\omega}E\cdot\nabla\times(\mu^{-1}\nabla\times E)dv\\&=-\frac{j}{\omega}\iint_S E\times(\mu^{-1}\nabla\times E)\cdot ndS.\end{aligned} \quad (31)$$

If $\sigma=0$ and E is real, (31) reduces to Berk's (6). If $\sigma\neq0$, (31) gives a complex value of ω. The real part is the resonant frequency and the imaginary part is simply related to the bandwidth.

To derive a formula based on an assumed distribution of E and H we take the source distribution as a volume distribution of electric dipoles J and magnetic dipoles K plus the surface distribution of magnetic dipoles K_s given by (29). In this case J is given by (26) and K is given by

$$K=-j\omega\mu H-\nabla\times E. \quad (32)$$

Substitution in (30) gives now

$$\begin{aligned}0&=\iiint_V(J\cdot E-K\cdot H)dv-\iint_S K_s\cdot HdS\\&=\iiint_V[-E(j\omega\epsilon+\sigma)E+E\cdot\nabla\times H\\&\qquad+H\cdot\nabla\times E+j\omega H\mu H]dv\\&\quad-\iint_S E\times H\cdot ndS.\end{aligned} \quad (33)$$

If E is pure real, H is pure imaginary and $\sigma=0$, (33) reduces to Berk's (10).

The corresponding formulas for the propagation constant of a waveguide follow from (30) in like manner, except that the formula for the reaction is different.[13,14] Starting from an assumed distribution of E and H over the cross section of the waveguide, (30) gives

$$\begin{aligned}0&=\iint_\Sigma(\mathcal{J}_s\mathcal{E}+\mathcal{K}_s\mathcal{H})dS+\oint_C\mathcal{K}_{ss}\mathcal{H}dl\\&=\iint_\Sigma[-\mathcal{E}_s(j\omega\epsilon+\sigma)\mathcal{E}-\mathcal{H}_sj\omega\mu\mathcal{H}\\&\qquad+\mathcal{E}_s\nabla\times\mathcal{H}-\mathcal{H}_s\nabla\times\mathcal{E}]dS\\&\quad+\oint_C\mathcal{H}_s(n\times\xi)dl\\&\quad+j\gamma\iint_\Sigma[\mathcal{E}_s(\hat{z}\times\mathcal{H})-\mathcal{H}_s(\hat{z}+\mathcal{E})]dS\end{aligned} \quad (34)$$

[12] Private communication.

[13] V. H. Rumsey, "Reaction concept in electromagnetic theory," *Phys. Rev.*, vol. 94, pp. 1483–1491; June, 1954, and vol. 95, p. 1705; September, 1954.

[14] V. H. Rumsey, "Travelling wave slot antennas," *J. Appl Phys.*, vol. 24, pp. 1358–1365; November, 1953.

where

$$E = e^{-i\gamma z}\mathcal{E}(xy) \text{ etc.}$$

s represents the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

n = unit outward normal to the waveguide wall, Σ = cross section of waveguide, C = cross section of wall, and $\hat{z}$ = unit vector in the z direction. Eq. (34) corresponds to Berk's (14). Equations corresponding to Berk's (15) and (16) can be derived in a similar manner.

Author's Reply

I was pleased to see that some of the formulas appearing in "Variational Principles for Electromagnetic Resonators and Waveguides" (VPERW) can be obtained from V. H. Rumsey's "reaction concept." (See Referee's Comments.) In this connection the following remarks are in order.

Professor Rumsey's formulas are variational provided the dielectric constant and permeability are, at most, symmetric tensors. This means that isotropic materials and materials with crystalline anisotropy, lossy or nonlossy, are allowed while gyrotropic media such as magnetized ferrites and magneto-ionic media are excluded.

The formulas of (VPERW), on the other hand, are restricted to nonlossy substances but they are applicable to gyrotropic media or media, in general, whose dielectric constant and permeability are hermitian tensors. Thus each set of formulas has its own sphere of applicability. If the complex conjugate sign is dropped the formulas of (VPERW) still remain variational provided their application is restricted to those cases where the expressions derived from the "reaction concept" are applicable.

Nonlinearity of Microwave Ferrite Media*

N. G. SAKIOTIS†, H. N. CHAIT†, AND M. L. KALES†

Summary—**Existing theories of propagation in magnetized ferrite media predict propagation constants which are independent of the rf field strength only if a number of restrictive conditions are satisfied. In general, however, the propagation constants are functions of the rf field strength and nonlinear propagation is to be expected. One of the conditions for linearity is that the rf magnetic field intensity be small compared to the static magnetizing field intensity. This condition can be violated when the peak power level of the wave incident on the ferrite medium is sufficiently high and indeed, it has been observed by the authors that ferrite loaded waveguides can become nonlinear at peak power levels as low as 1 kw.**

Results are presented of a study of the behavior of ferrite loaded waveguides at 9,400 mc over input peak power levels from 0.1 kw to 100 kw which includes power levels commonly encountered in radar applications. The dependence upon input power level of the ferrite losses, phase shift, and rotation of the plane of polarization is described for a number of ferrite materials. The effect upon the degree of nonlinearity of the intensity of the static magnetizing field is discussed as well as the dependence upon the dimensions of the ferrite.

POLDER[1] and others have developed a theory dealing with the propagation of electro-magnetic waves in a magnetized ferrite medium. According to this theory, the propagation will be linear with respect to the rf field strength only if a number of restrictive conditions are fulfilled. Under the most general conditions, however, nonlinear propagation is to be expected. Among the assumptions which lead to a linear theory is one which requires that the rf magnetic field should be small compared to the static magnetic field. Another is that the damping constant is independent of power level. The first of these assumptions will clearly be violated if the peak power level is sufficiently high. The theory in its present form does not indicate how large the rf magnetic field must become before an appreciable departure from linearity can be observed. For this reason the authors have conducted tests to determine if any nonlinearity is evident at power levels which are normally encountered in radar applications. Actually it turns out that the propagation becomes nonlinear at relatively low power levels.

This situation is very much of practical interest since the authors have observed that the nonlinear behavior can occur at peak power levels as low as 0.7 kw and the absorption loss can become large enough at peak power levels less than 100 kw so as to make ferrite devices unusable in many radar applications, even though these devices appeared low-loss when tested at klystron power levels.

An example of the increase in absorption can be seen in Fig. 1. Here we have a ferrite cylinder, 0.0225 inch in diameter, 6 inches long, mounted on the axis of an X-band circular waveguide and magnetized by a static magnetic field in the direction of the axis. The wave is circularly polarized and the loss due to absorption of energy by the ferrite is plotted, for each of the two senses of polarization, as a function of the peak power incident upon the ferrite.

The positive sense of circular polarization refers to that polarization whose configuration rotates, with

* Manuscript received by PGAP August 17, 1955.
† Microwave Antennas and Components Branch, Electronics Div., Naval Res. Lab., Washington, D. C.
[1] D. Polder, "On the theory of ferromagnetic resonance," *Phil. Mag.*, vol. 40, pp. 99–115; January, 1949

time, in the direction that would advance a right-handed screw in the direction of the applied magnetizing field.

It can be seen from the curve for the wave circularly polarized in the negative sense that as the peak power level is increased, the absorption loss remains constant at 1.2 db, which is the absorption loss observed at klystron power levels, until the peak input power reaches about 1 kw. As the input power is increased beyond this level the fraction of the input power absorbed, in db, rises, reaching 18 db at the 100 kw level. The wave polarized in the positive sense typically exhibits the same behavior except that the power level at which the absorption begins to increase is higher, approximately 10 kw, and the rate of increase with input power level is smaller.

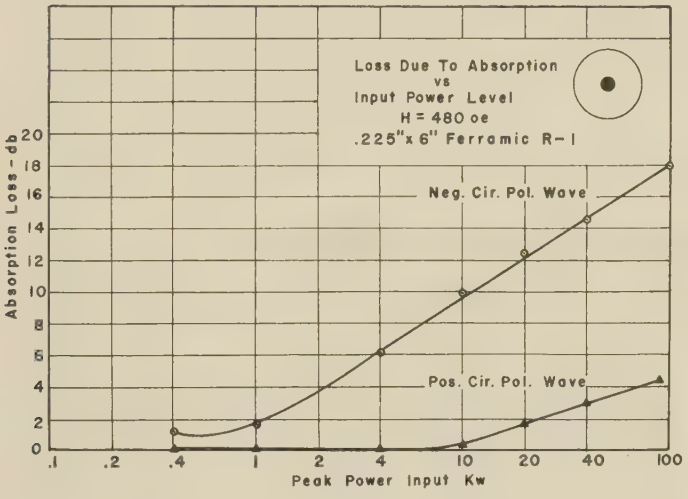


Fig. 1

The linear theory for a ferrite medium magnetized along the direction of propagation predicts a different effective permeability for each of the two senses of circular polarization. This difference in permeability gives rise to a difference in the rf field strength in the ferrite for the two senses of circular polarization. Therefore the nonlinear behavior should be different for the two cases. It is clear from the curves that the nonlinear effects do indeed depend on the sense of circular polarization.

It is interesting that the Faraday rotation, which is due to the unequal phase velocities of the two senses of circular polarization, has been observed to be independent of the power level up to a peak power level of 100 kw for the materials and dimensions we have tested.

During this experiment the magnetizing field was kept constant at 480 oersteds which is the magnitude of applied field necessary to produce the maximum amount of Faraday rotation for the ferrite samples that will be discussed.

For the data to be discussed, the power source was a magnetron, operating at 9,400 mc, modulated by 1 microsecond pulses at a repetition rate of 1,000 cycles. Thus the average power is one-thousandth of the peak power.

Since the propagation characteristics of a ferrite medium are also temperature sensitive, care had to be taken to insure that the observed variation of attenuation with power level was not due to a rise in temperature of the ferrite, since the ferrite samples had to dissipate 100 watts at times. This was checked in a number of ways. One such check involved a comparison of the high power data with temperature data. The attenuation and rotation were first measured as a function of power level. Then with the power level held at a low constant value the attenuation and rotation were measured as a function of temperature. It was found that when the temperature was raised to a level high enough to obtain losses comparable to those observed at high power levels, the change in rotation was very large compared to the change in rotation resulting from an increase in power level. If the heating of the ferrite at high power had been the primary cause of the observed changes in properties, one would expect the changes in rotation to be comparable. For example, it was found that for a ferrite rod 0.225 inch in diameter and 2 inches long the loss increased from 0.5 db to 3 db, and the rotation changed at most 5° when the peak power level was increased from 0.4 kw to 10 kw. In order to get a 3 db loss at the low power level the temperature had to be raised to more than 550°F. At this temperature the change in rotation was 120°.

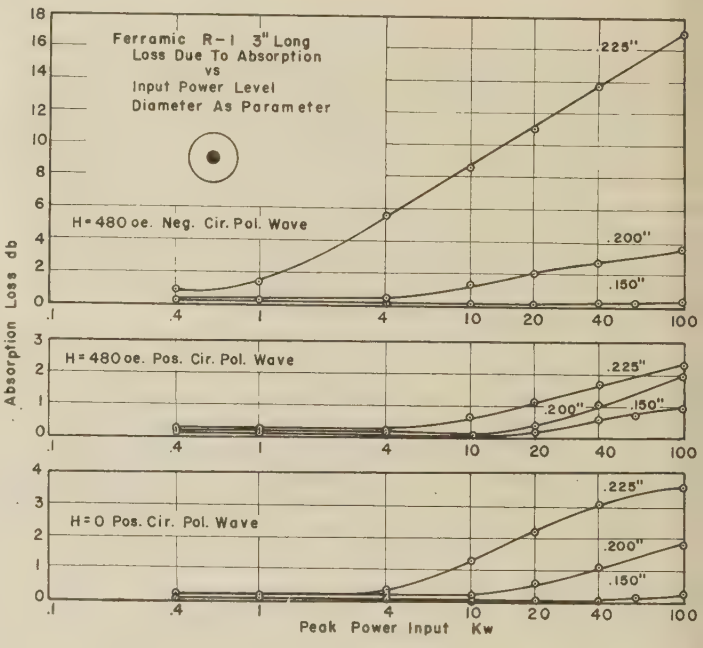


Fig. 2

For a given input power level, the magnitude of the rf magnetic field intensity in the ferrite will also be a function of the ferrite diameter. As the diameter is decreased, the magnitude of the rf field intensity decreases. Thus one would expect the power level at which the ferrite becomes nonlinear to be dependent upon the ferrite cylinder diameter. Fig. 2 shows the observed dependence on diameter for both senses of polarization

when the ferrite is magnetized by an applied field of 480 oersteds and also for the unmagnetized case. It is seen that as the diameter of the ferrite cylinder decreases, the amount of rf power required to bring about the nonlinear behavior of the absorption loss increases. This is especially evident in the case of the negative circular polarization. The 0.150 diameter cylinder maintains its low-power absorption characteristics up to 60 kw while the 0.225 diameter cylinder becomes nonlinear in the vicinity of 0.7 kw and at 100 kw the absorption loss has increased by 16 db.

The tendency toward more linear propagation as the diameter decreases is also exhibited in the curves for the positive sense of polarization and in the unmagnetized case.

Another factor which affects the nonlinear behavior of the ferrite is the intensity of the applied static magnetizing field. In fact, if the ferrite diameter is small, it can be shown that the rf magnetic intensity in the unmagnetized ferrite decreases upon application of a magnetizing field, for the negative sense of polarization, and increases, for the positive sense. It might therefore be expected that, for rods of sufficiently small diameter, the input power level necessary for the incidence of the nonlinear behavior is higher for waves circularly polarized in the negative sense than it is for waves circularly polarized in the positive sense. This behavior can be seen in the curves for the smallest diameter rod.

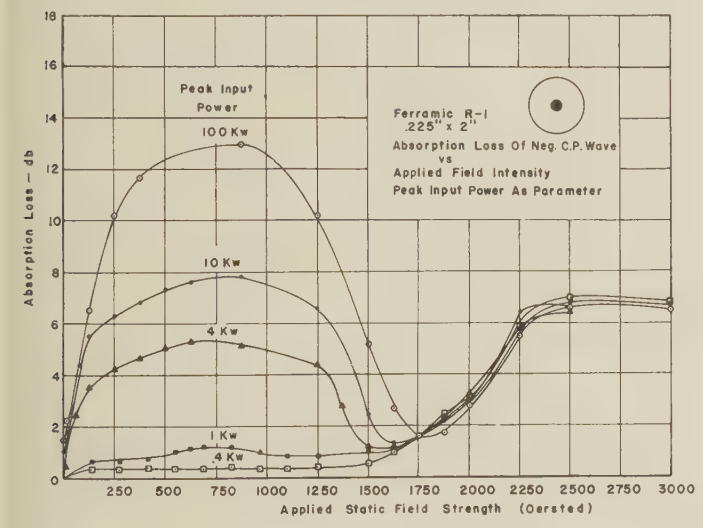


Fig. 3

The effect of the applied static magnetizing field in the case of a larger diameter rod is shown in Fig. 3. Each of the curves represents the absorption loss as a function of the applied magnetizing field strength with the incident peak rf power level kept constant. It can be seen from these curves that the change in absorption loss between the lowest and highest power levels increases with increasing applied field, reaching a maximum at approximately 750 oersteds, then decreases until at an applied field of 1,750 oersteds it disappears. Above this value of applied field strength there appears to be a small amount of change, the absorption loss decreasing with an increase in the rf power level. Thus the effect of nonlinearity on absorption seem to be greatest in the region of applied field strength below 1,000 oersteds, which is the region most commonly used in the design of microwave ferrite phase shifters and rotators.

Fig. 4 shows the same type of data taken for the wave circularly polarized in the positive sense. Qualitatively the behavior is much the same as for waves circularly polarized in the negative sense for applied fields below 2,000 oersteds.

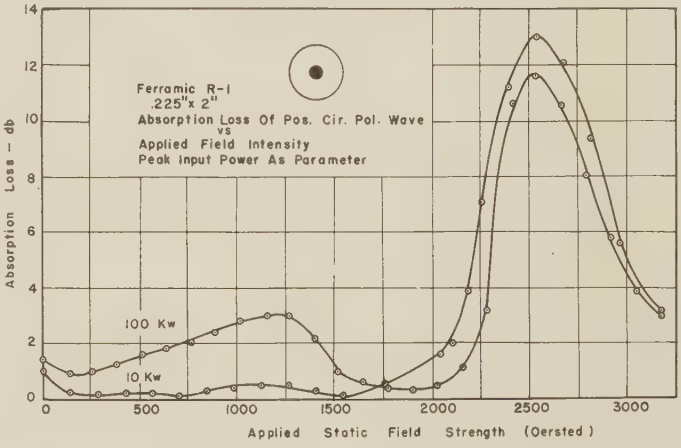


Fig. 4

Ferromagnetic resonance absorption occurs for this sense of circular polarization, and we see the characteristic resonance loss peak at the higher power level is smaller and narrower. This decrease in absorption at the resonance field strength as the rf field intensity is increased has also been observed by Bloembergen[2] and his co-workers using cavity techniques.

An estimate of the dependence of the effective attenuation constant upon the microwave power level can be obtained by studying the absorption loss as a function of the length of the sample, maintaining the sample diameter and the magnetizing field strength constant.

In Fig. 5 are shown curves of absorption loss vs ferrite cylinder length, for the negative sense of circular polarization, the diameter held fixed at 0.225 inch. For each curve the peak power incident on the ferrite cylinder is constant. Looking at the 20 kw curve, the absorption in db per unit length is almost constant up to a length of about 2 inches. Beyond this length, there is an interval in which the absorption loss in db per unit length decreases rapidly with length. As the length of the sample is further increased, or if the incident power level is decreased, the absorption per unit length is again constant. Since the power level has decreased by about 8.5 db in traversing the first two inches of the ferrite, the power level at that point is about 3 kw. Therefore for power levels between 20 kw and 3 kw it

[2] N. Bloembergen and S. Wang, "Relaxation effects in para- and ferromagnetic resonance," *Phys. Rev.*, vol. 93, pp. 72–83; January, 1954.

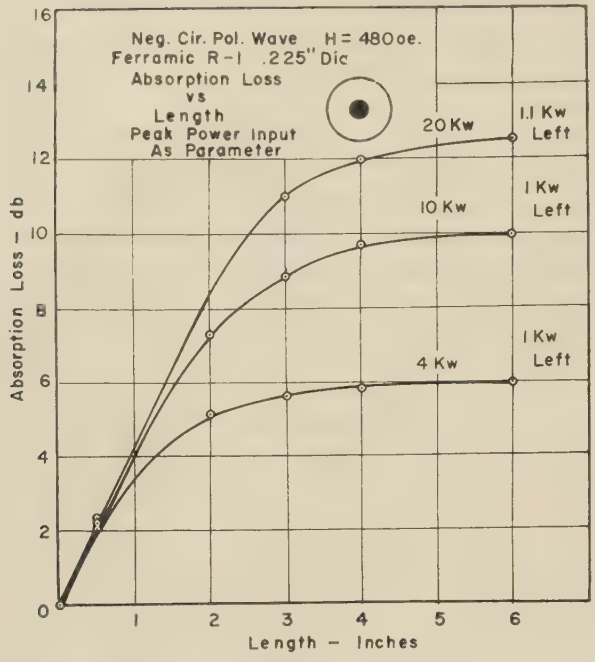


Fig. 5

appears that the effective attenuation constant does not change very much with rf power level. It can be seen from the curves of Fig. 5 that the effective attenuation constant is also constant at a small value for power levels below 1 kw. Between 1 kw and 3 kw there appears a large change in the effective attenuation constant.

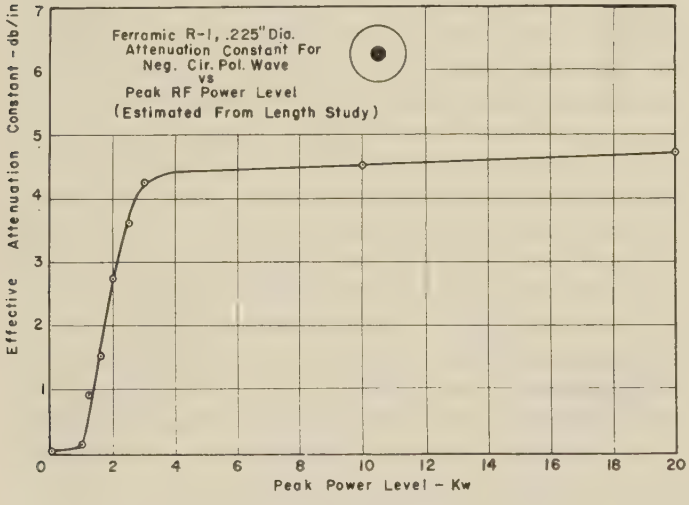


Fig. 6

Fig. 6 shows a curve of the effective attenuation constant for the 0.225-inch diameter ferrite rod computed from the slopes of the absorption vs length curves. The transition region from the low loss to the high loss behavior between 1 kw and 3 kw can be seen quite clearly.

Thus we can visualize what happened in the rod as follows. If the incident power is more than 3 kw peak, the attenuation constant of the material is at a high value over that length of the material which is necessary to drop the power level to 3 kw. This is followed by a relatively short length of material in which the attenuation constant is decreasing rapidly. After the power level had dropped to about 0.7 kw peak, the attenuation constant in the remainder of the length is again appreciably constant at a relatively small value. Therefore as the input power level is changed, keeping the ferrite dimensions the same, the percentage of the total length which has a high attenuation constant changes, resulting in a total absorption which varies with input power levels in the manner of the experimental curves.

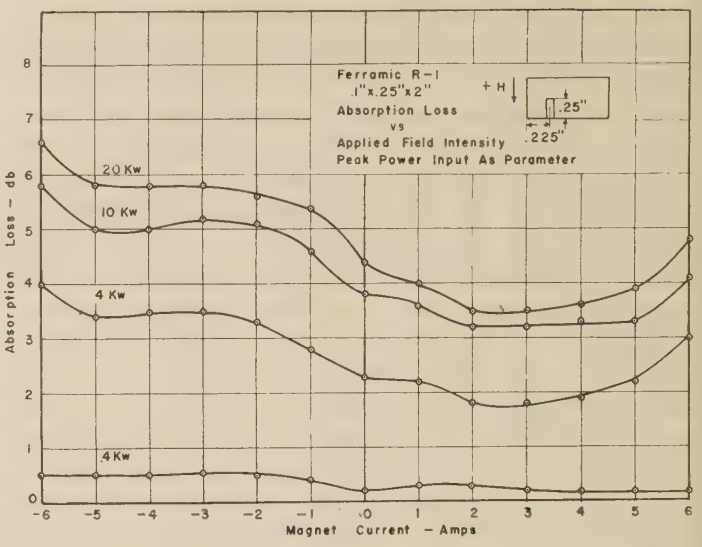


Fig. 7

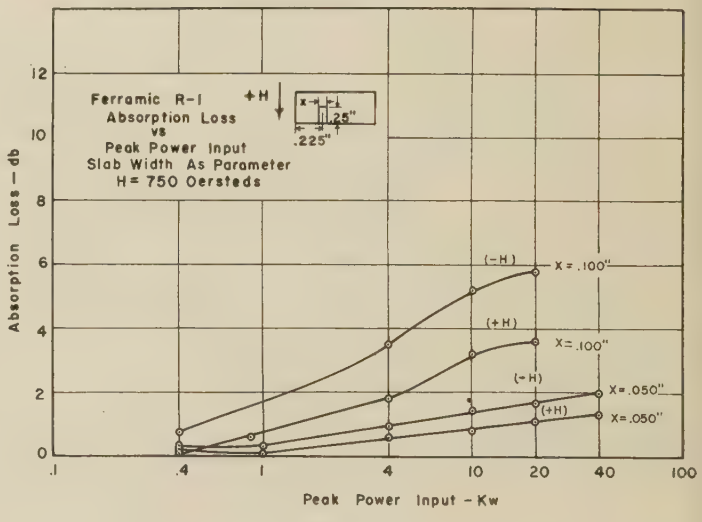


Fig. 8

Qualitatively, much the same nonlinear behavior is exhibited by ferrites in rectangular waveguide. Fig. 7 shows the absorption loss of a ferrite slab in X-band rectangular waveguide plotted as a function of the applied magnetizing field for a number of input peak power levels. The range of applied field strength represented here is from $-1{,}500$ oersteds to $+1{,}500$ oersteds. The increase in absorption loss with increasing power level is quite evident. As in the case of the circular waveguide configuration, it was found that as the slab is made thinner, the incidence of the nonlinear behavior

occurs at higher power levels. This is shown in Fig. 8. Here the applied magnetizing field is held constant at 750 oersteds and the loss is plotted as a function of the input peak power level for both directions of magnetization. It is seen that the curves for the thinner piece begin to rise at a higher value of incident power.

During the past few years, considerable effort has been made to produce better low-loss ferrite materials for microwave applications. Much of the evaluation of new materials has been made at low klystron power levels. The presence of the nonlinear effects makes it necessary to produce materials that are not only low-loss at very small power levels but low-loss at those power levels at which they will ultimately be used.

Similarly, any design of a ferrite component should be tested at the power level and with the waveguide geometry of the final product since the rf field intensity will in general depend upon the waveguide-ferrite geometry as well as the power level of the incident wave.

Although ferrites exhibiting the nonlinear behavior which has been described would not be suitable for many applications, it may well turn out that new applications which make use of the nonlinear behavior will be devised. This possibility is under study.

Diffraction of Electromagnetic Waves Caused by Apertures in Absorbing Plane Screens*

HANS E. J. NEUGEBAUER†

Summary—**The perturbation field caused by holes in a plane, infinitely thin screen of arbitrary material on which electromagnetic radiation is incident, can be split up in two fields, one of which is symmetrical in the tangential components of the electric, the other in those of the magnetic vector. Both fields satisfy equations within the holes which are generalizations of Bethe's conditions for the perfectly conducting screen and which can be used to derive approximate solutions of Kirchhoff type. The solution for a screen which absorbs practically all of the incident radiation is essentially different from Kirchhoff's result for scalar radiation incident on a so-called perfectly absorbing screen.**

Short Review of the Theory for the Perfectly Conducting Screen

THIS PAPER is concerned with relations for the electromagnetic diffraction field caused by holes in a screen of any material, which are generalizations of Bethe's[1] conditions for the perfectly conducting screen.

Bethe's conditions constitute a statement about the perturbation field and, consequently, about the total field in the aperture. Being rigorous for the ideal case of a perfectly reflecting screen of zero thickness, they can be used to derive an integral equation for the diffraction field.[2] Another application is to the problem of calculating the perturbation field which can be formulated as a mixed boundary value problem.[3] Rigorous solutions are known for the infinite half-plane[4] and for a circular aperture[5] in a perfectly reflecting screen. Yet, the greater value of these conditions seems to lie in the fact that they can be used as basis for approximate solutions.[6] Such solutions which can also be derived by different methods[7–9] are of a similar nature as Kirchhoff's solutions of scalar diffraction problems and can be obtained for various shapes of the holes. An extensive experimental study under G. A. Woonton has shown that theory and experiment are in good agreement providing the apertures are larger than a wave length.[10] Therefore, this approach seems to have definite advantages.

The generalized Bethe's conditions to be derived in this paper hold rigorously for screens of any material providing screen thickness is small compared to free space wave length. They can be used, in a very similar way to the one mentioned above, for obtaining approximate solutions for the perturbation fields caused by holes in a screen.

* Manuscript received by the PGAP, August 4, 1955. Revised manuscript received January 12, 1956. The basic ideas of this paper are contained in this author's Technical Report No. 25, November 1952, Eaton Electronics Res. Lab., McGill University, Montreal, Canada under USAF Contract AF 19(122)-81 through Cambridge Res. Ctr.

† Formerly, Adalia Ltd., Montreal; now, Res. Labs., RCA Victor Co., Ltd., Montreal.

[1] H. A. Bethe, "Theory of diffraction by small holes," *Phys. Rev.*, vol. 66, pp. 163–182; October, 1944.

[2] B. B. Baker and E. T. Copson, "The Mathematical Theory of Huygens' Principle," Clarendon Press, Oxford, Eng., p. 164; 1950.

[3] G. Bekefi, "Diffraction of electromagnetic waves by an aperture in an infinite screen," *J. Appl. Phys.*, vol. 23, p. 1403; December, 1952, "Diffraction of electromagnetic waves by an aperture in a large screen," vol. 24, 1123; September, 1953.

[4] A. Sommerfeld, "Ueber die ausbreitung der wellen in der drahtlosen telegraphie," *Ann. Physik*, vol. 28, p. 665; March, 1909; vol. 81, p. 1135; December, 1926.

[5] J. Meixner and W. Andrejewski, "Strenge theorie der beugung ebener elektromagnetischer wellen an der vollkommen leitenden kreisscheibe und an der kreisfoermigen oeffnung im vollkommen leitenden ebenen schirm," *Ann. Physik.* (6), vol. 7, p. 157; April, 1950. W. Andrejewski, "Strenge theorie der beugung ebener elektromagnetischer wellen an der vollkommen leitenden kreisscheibe und an der kreisfoermigen oeffnung im vollkommen leitenden ebenen schirm," *Naturwiss.*, vol. 38, p. 406; September, 1951.

[6] S. Silver and M. J. Ehrlich, Antenna Lab. Reports No. 181 (1951) and No. 185 (1952), Dept. of Engrg., University of California.

[7] H. Severin, "Beugung elektromagnetischer zentimeterwellen an metallischen blenden," *Z. Naturforsch.*, vol. 1, p. 487; September, 1946.

[8] H. E. J. Neugebauer, Eaton Electronics Res. Lab. Report No. b5 on contract to the USAF, Cambridge Res. Ctr. (1952); "A new method of solving diffraction problems," *J. Appl. Phys.*, vol. 23, p. 1406; December, 1952.

[9] S. J. Buchsbaum, Eaton Electronics Res. Lab. Report No. b9 on contract to the USAF, Cambridge Res. Ctr., 1953.

[10] S. J. Buchsbaum, A. R. Milne, D. C. Hogg, G. Bekefi, and G. A. Woonton, "Microwave diffraction by apertures of various shapes," *J. Appl. Phys.*, vol. 26, p. 706; June, 1955.

Some introductory remarks on the theory for the perfectly reflecting screen may be given before the general case will be discussed in the succeeding sections.

If electromagnetic radiation of harmonic time dependence, $\exp(j\omega t)$, is incident from the left side, $z<0$, on a plane and perfectly conducting, infinitely thin screen with holes of any shape and size, the total field before and behind the screen can be described as the sum of the incident, reflected and perturbation field. Since incident and reflected field, $\vec{E}^i$, $\vec{H}^i$ and $\vec{E}^r$, $\vec{H}^r$, are those which would exist if the holes were closed, they are zero behind the screen, $z>0$. If the holes are open the perturbation field, $\vec{E}^p$, $\vec{H}^p$, is added to the incident and reflected field, and the tangential components $E_x{}^p$, $E_y{}^p$ of the electric vector as well as the normal component $H_z{}^p$ of the magnetic vector are shown to be symmetrical with regard to the plane of the screen, $z=0$. Other components $E_z{}^p$, $H_x{}^p$, $H_y{}^p$ are antisymmetrical.

These symmetry relations can be derived from Stratton and Chu's vector Green's theorem.[11] The simplest way, however, to establish them is to add fictitious sources lying symmetrical to the real sources with respect to the screen and radiating with opposite phase.[12] The incident and reflected fields of the fictitious sources are the continuations, through the plane $z=0$, of the reflected and incident fields respectively of the real sources so that both types of sources together do not produce any perturbation field.

Therefore, the perturbation field of the fictitious sources at any point is opposite that of the real sources at the same point. In addition, due to the phase relation mentioned above, the perturbation field of the fictitious sources is antisymmetrical to that of the real sources in the E_x, E_y, and H_z components, but symmetrical in the E_z, H_x, and H_y components. Hence follow the symmetry properties of the perturbation field.

The further condition that the total field, $\vec{E}^i+\vec{E}^r+\vec{E}^p$, $\vec{H}^i+\vec{H}^r+\vec{H}^p$, must be continuous in the holes, together with its first derivatives, leads to simple equations relating the perturbation field to the incident and reflected field. They are known as Bethe's[1] conditions and are a special case of (5).

Symmetrical and Antisymmetrical Perturbation Field

Let it now be assumed that the screen consists of any material and that its thickness is negligible compared to free space wave length. The further assumption that the screen may be opaque is made only to simplify the formulas since this theory can be carried through in essentially the same way for a transparent screen. The required additions for a transparent screen will be mentioned.

The total field is the sum of the incident, reflected and perturbation field.

$$\begin{cases} \vec{E} = \vec{E}^i + \vec{E}^r + \vec{E}^p \\ \vec{H} = \vec{H}^i + \vec{H}^r + \vec{H}^p. \end{cases} \tag{1}$$

Incident and reflected fields may be known.

Only the coordinate z of the field point will be written in the following equations while x and y will be suppressed as they are always the same for all field functions occurring in the same equation.

The perturbation field can be split up into the sum of two fields, $\vec{E}^s$, $\vec{H}^s$ and $\vec{E}^a$, $\vec{H}^a$, by means of the following:

$$\begin{cases} * \quad E_{x,y}{}^s(z) = \tfrac{1}{2}[E_{x,y}(z) + E_{x,y}(-z)] \\ \quad E_z{}^s(z) = \tfrac{1}{2}[E_z(z) - E_z(-z)] \\ \quad H_{x,y}{}^s(z) = \tfrac{1}{2}[H_{x,y}(z) - H_{x,y}(-z)] \\ * \quad H_z{}^s(z) = \tfrac{1}{2}[H_z(z) + H_z(-z)] \\ \quad E_{x,y}{}^a(z) = \tfrac{1}{2}[E_{x,y}(z) - E_{x,y}(-z)] \\ * \quad E_z{}^a(z) = \tfrac{1}{2}[E_z(z) + E_z(-z)] \\ * \quad H_{x,y}{}^a(z) = \tfrac{1}{2}[H_{x,y}(z) + H_{x,y}(-z)] \\ \quad H_z{}^a(z) = \tfrac{1}{2}[H_z(z) - H_z(-z)]. \end{cases} \tag{2}$$

It can easily be verified that

a) $\vec{E}^p=\vec{E}^s+\vec{E}^a$ and $\vec{H}^p=\vec{H}^s+\vec{H}^a$;

b) both partial fields, $\vec{E}^s$, $\vec{H}^s$ as well as $\vec{E}^a$, $\vec{H}^a$, satisfy Maxwell's equations in free space;

c) the field components marked with an asterisk in (2) are symmetrical, the others are antisymmetrical with respect to the plane of the screen, $z=0$.

As regards the nomenclature, the tangential field components of the electric vector are considered representative. Hence, the partial field $\vec{E}^s$, $\vec{H}^s$ will be called the symmetrical, $\vec{E}^a$, $\vec{H}^a$ the antisymmetrical perturbation field.

For future applications it may be noted that the derivative, with respect to z, of a symmetrical field component is antisymmetrical and vice versa.

It may also be noted that Kirchhoff's expression for the perturbation field in the case of scalar waves is known to be the sum of a symmetrical and an antisymmetrical function.[13]

Generalized Bethe's Conditions

Applying the continuity condition of the field within an aperture to the x-component of the electric vector,

$$E_x{}^i(-0) + E_x{}^r(-0) + E_x{}^s(-0) + E_x{}^a(-0) = E_x{}^i(+0) + E_x{}^r(+0) + E_x{}^s(+0) + E_x{}^a(+0). \tag{3}$$

Because of

$$E_x{}^i(+0) = E_x{}^r(+0) = 0, \quad E_x{}^s(-0) = E_x{}^s(+0)$$

and

$$-E_x{}^a(-0) = E_x{}^a(+0),$$

[11] J. A. Stratton and L. J. Chu, "Diffraction theory of electromagnetic waves," *Phys. Rev.*, vol. 56, pp. 99–107; July, 1939.

[12] B. B. Baker and E. T. Copson, *loc. cit.*, pp. 159–160, 163–164.

[13] C. J. Bouwkamp, "Diffraction theory," *Phys. Soc., Prog. Rep. in Phys. XVII*, p. 52; 1954.

(3) reduces to

$$E_x^a(+0) = \tfrac{1}{2}[E_x^i(-0) + E_x^r(-0)]. \quad (4)$$

Correspondingly,

$$H_x^i(-0) + H_x^r(-0) + H_x^s(-0) + H_x^a(-0) = H_x^i(+0) + H_x^r(+0) + H_x^s(+0) + H_x^a(+0)$$

which reduces, because of $H_x^i(+0)=H_x^r(+0)=0$, $H_x^s(-0)=-H_x^s(+0)$ and $H_x^a(-0)=+H_x^a(+0)$, to

$$H_x^s(+0) = \tfrac{1}{2}[H_x^i(-0) + H_x^r(-0)]. \quad (5)$$

All the symmetrical and antisymmetrical field components satisfy similar equations to (4) and (5) respectively. This holds true also for the derivatives with respect to z.

In the case of a perfectly conducting screen

$$\vec{E}^a = \vec{H}^a \equiv 0$$

$$E_{x,y}^r(-0) = -E_{x,y}^i(-0), \quad E_z^r(-0) = E_z^i(-0)$$

$$H_{x,y}^r(-0) = H_{x,y}^i(-0), \quad H_z^r(-0) = -H_z^i(-0)$$

and (5) simplifies to one of Bethe's conditions,

$$H_x^s(+0) = H_x^i(-0).$$

For a transparent screen the unperturbed field consists of the incident and reflected field,

$$\vec{E}^i, \vec{H}^i \quad \text{and} \quad \vec{E}^r, \vec{H}^r,$$

in front of the screen, $z>0$, and of the transmitted field,

$$\vec{E}^t, \vec{H}^t,$$

behind the screen, $z>0$. Hence, the following changes must be made if the screen is transparent:

$\vec{E}^t$ and $\vec{H}^t$ respectively are added to the right sides of (1).

$E_x^t(+0)$ is added to the right side of (3) and is subtracted in the square bracket of (4). Similarly, $H_x^t(+0)$ is subtracted in the square bracket of (5).

It may be noted that the derivations of this section are rigorous providing screen thickness and multiple scattering between sources and holes can be neglected.[4,5]

Approximate Solutions

Generalized Bethe's conditions (4) and (5) can be used to calculate approximate solutions of Kirchhoff type for the diffraction field in cases where the reflected field can be determined. The general idea is to assume the field in the shadow region near the screen as negligibly small so that the tangential magnetic components of the symmetrical field and the tangential electrical components of the antisymmetrical field can be determined by means of Green's theorem,

$$H_{x,y}^s(P) = \frac{1}{2\pi}\int_{\text{aperture}} H_{x,y}^s(Q)\ \frac{\partial}{\partial z_Q}\left[\frac{\exp(-jkr_{QP})}{r_{QP}}\right] ds_Q \quad (6)$$

$$E_{x,y}^a(P) = \frac{1}{2\pi}\int_{\text{aperture}} E_{x,y}^a(Q)\ \frac{\partial}{\partial z_Q}\left[\frac{\exp(-jkr_{QP})}{r_{QP}}\right] ds_Q \quad (7)$$

where P is a field point behind the screen, $z_P>0$, Q is a point within the aperture, r_{QP} is the distance between Q and P. Reference is made to Fig. 1. Field components $E_{x,y}^a(Q)$ and $H_{x,y}^s(Q)$ are found from (4) and (5).

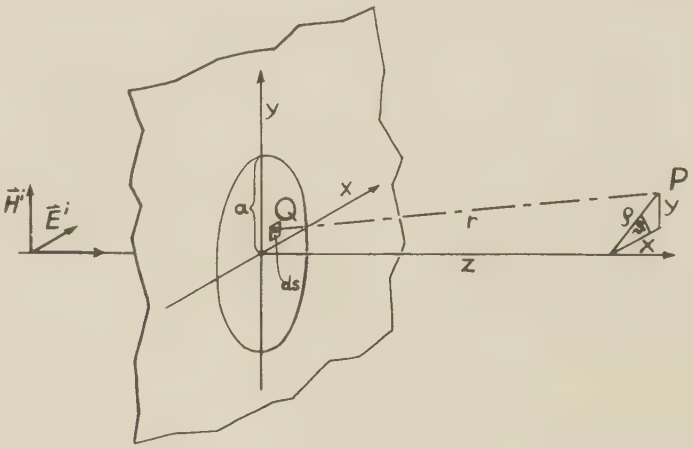


Fig. 1

Although more general cases can be treated, the equations for a plane wave normally incident on a circular hole may be given here as an example. Let the incident field, without the time dependent factor, be

$$\begin{cases} E_x^i = E_0 \exp(-jkz), & E_y^i = E_z^i = 0 \\ H_y^i = \dfrac{kE_0}{\omega\mu}\exp(-jkz), & H_x^i = H_z^i = 0. \end{cases} \quad (8)$$

Then, the reflected field is

$$\begin{cases} E_x^r = -ZE_0 \exp(+jkz), & E_y^r = E_z^r = 0 \\ H_y^r = \dfrac{ZkE_0}{\omega\mu}\exp(+jkz), & H_y^r = H_y^r = 0, \end{cases} \quad (9)$$

where Z is the reflection coefficient.

From (4) to (9) it follows

$$H_y^s = \frac{kE_0(1+Z)}{4\pi\omega\mu}\int_{\text{aperture}} \frac{\partial}{\partial z}\left[\frac{\exp(-jkr)}{r}\right] ds \quad (10)$$

$$E_x^a = \frac{E_0(1-Z)}{4\pi}\int_{\text{aperture}} \frac{\partial}{\partial z}\left[\frac{\exp(-jkr)}{r}\right] ds. \quad (11)$$

Apart from the factor $\frac{1}{2}(1+Z)$ (10) is the expression yielding the total perturbation field for the perfectly reflecting screen. In addition, both factors $\frac{1}{2}(1+Z)$ and $\frac{1}{2}(1-Z)$ being disregarded, the antisymmetrical field, (11), can be derived from the symmetrical, (10), by proper interchanging of E and H. Hence, the total perturbation field is a linear superposition of the perturbation fields caused by a perfect electrical and a perfect magnetic conductor.

One limiting case of special interest is constituted by the perfectly reflecting screen, $Z=1$. The other limiting

case is constituted by a screen that absorbs so much of the incident energy that the reflected and transmitted fields can be neglected, $Z=0$. Although it is difficult to make such a screen whose thickness is negligible compared to free space wave length, it is of interest to investigate the nature of the field in this limiting case. The screen would be perfectly absorbing in the sense of Kirchhoff's theory as practically no energy is reflected from, nor transmitted through, the closed screen, although it may not be perfectly absorbing for grazing incidence. Since, however, we confine these derivations to the case of normal incidence and since we consider the perturbation field near the screen as negligibly small, this shortcoming of realizable screens should not weigh too heavily.

A screen that transmits or reflects only negligible amounts of normally incident radiation may be called "zero reflection screen." It is characterized by $Z=0$. The perturbation field is the sum of the two fields given by

$$H_y^s = \frac{E_0 k}{4\pi\omega\mu}\int \frac{\partial}{\partial z}\left[\frac{\exp(-jkr)}{r}\right]ds \qquad (12)$$

and

$$E_x^s = \frac{E_0}{4\pi}\int \frac{\partial}{\partial z}\left[\frac{\exp(-jkr)}{r}\right]ds. \qquad (13)$$

This result is essentially different from Kirchhoff's solution for the perfectly absorbing screen in the scalar case where also two integrals must be calculated, one of which is

$$\int u\frac{\partial}{\partial z}\left[\frac{\exp(-jkr)}{r}\right]ds$$

as above, but the other one is

$$\int \frac{\exp(-jkr)}{r}\left(\frac{\partial u}{\partial z}\right)ds.$$

Eqs. (12) and (13) are in agreement with an earlier solution derived by this author[14] with the help of the Hertz vector Π for the symmetrical, and the Π-vector for the antisymmetrical field. However, they are at variance with an earlier solution[8] based on Π alone and which does not satisfy the generalized Bethe's conditions unknown at that time.

Since the total perturbation field is a superposition of two fields, one caused by a perfect electrical, the other one by a perfect magnetic conductor, and since the latter follows from the former by interchanging of E and H, it is sufficient to evaluate only (10) for the case of a perfect conductor, $Z=1$.

From (10), because of

$$\frac{\partial}{\partial z_Q} = -\frac{\partial}{\partial z_P},$$

$$H_y^s(P) = -\frac{kE_0}{2\pi\omega\mu}\frac{\partial}{\partial z_P}\int_{\text{aperture}} \frac{\exp(-jkr)}{r}ds.$$

By means of Green's theorem the integral over the aperture can be transformed in an expression containing an integral over the cylindrical surface whose generatices are the rays, in the geometrical optics sense, through the rim of the hole. The integration along a generatrix can be performed leading to a contour integral of the form

$$\int_0^{\pi} \frac{a-\rho\cos\theta}{a^2+\rho^2-2a\rho\cos\theta} \cdot\exp\left[-jk(a^2+\rho^2-2a\rho\cos\theta+z^2)^{1/2}\right]d\theta \qquad (14)$$

where $\rho=+\sqrt{x^2+y^2}$ and $\tan\theta=y/x$. Up to this point the calculations are almost the same as those published by Rubinowicz.[15] For further evaluation the field points are confined to the space near the perpendicular axis through the center of the hole,

$$(\rho/a)^2 \ll 1.$$

With $a^2+z^2=A^2$ and terms of higher order in ρ/a being neglected, integral (14) becomes

$$\int_0^{2\pi} \frac{\exp(-jkA)}{a}\left(1+\frac{\rho}{a}\cos\theta\right)\exp\left(\frac{a\rho}{A}jk\cos\theta\right)d\theta$$

which can be solved by means of the Bessel functions J_0 and J_1 of first kind.

Finally the electric field is derived from the tangential magnetic components by means of

$$\vec{\nabla}\cdot\vec{H}=0 \quad \text{and} \quad \vec{E}=-\frac{j}{\omega\epsilon}\vec{\nabla}\times\vec{H}.$$

The result in a form suitable for experimental tests, shall be quoted for a screen of any material with a large circular aperture, $ka\gg 1$, and for field points behind the screen, $z>0$, which lie either in the H plane, $x=0$, or in the E plane, $y=0$, near the axis through the center of the hole, $\rho^2\ll a^2$.

a) *H-plane* $(x=0)$:

$$E_x^p = \frac{1+Z}{2}e_x + \frac{1-Z}{2}\frac{k}{\omega\epsilon}h_y'$$

$$E_y^p = 0$$

$$H_x^p = 0$$

$$H_y^p = \frac{1+Z}{2}h_y + \frac{1-Z}{2}\frac{k}{\omega\mu}e_x'$$

b) *E-plane* $(y=0)$:

$$E_x^p = \frac{1+Z}{2}e_x' + \frac{1-Z}{2}\frac{k}{\omega\epsilon}h_y$$

$$E_y^p = 0$$

$$H_x^p = 0$$

$$H_y^p = \frac{1+Z}{2}h_y' + \frac{1-Z}{2}\frac{k}{\omega\mu}e_x$$

[14] H. E. J. Neugebauer, Symposium on Microwave Optics, Montreal, Canada, Paper No. 51, p. 4; June, 1953.

[15] A. Rubinowicz, "Die Beugungswelle in der kirchhoffschen theorie der beugungserscheinungen," *Annalen der Physik* (4), vol. 53, pp. 257–278; August, 1917.

with

$$e_x = E_0 \left\{ \exp(-jkz) - \exp(-jkA) \cdot \left[J_0(kB) - \frac{a^2}{A^2} \frac{J_1(kB)}{kB} \right] \right\}$$

$$e_x' = E_0 \left\{ \exp(-jkz) - \exp(-jkA) \cdot \left[\left(1 - \frac{a^2}{A^2}\right) J_0(kB) + \frac{a^2}{A^2} \frac{J_1(kB)}{kB} \right] \right\}$$

$$h_y = h_y' = \frac{kE_0}{\omega\mu} \left\{ \exp(-jkz) - \frac{z}{A} \exp(-jkA) \cdot \left[J_0(kB) + j\left(1 + \frac{a^2}{A^2}\right) \frac{\rho}{a} J_1(kB) \right] \right\}$$

$$B = \frac{\rho a}{A} = \begin{cases} \dfrac{xa}{A} & \text{for the } H \text{ plane} \\ \dfrac{xa}{A} & \text{for the } E \text{ plane,} \end{cases}$$

$$k^2 = \omega^2 \epsilon \mu.$$

Further Applications

The theory briefly described in the preceding section has been shown to be in good agreement with experiments for different sizes and shapes of holes in a perfectly conducting screen.[10] Experiments on screens of zero reflection have been started by several workers. No results have been published yet, probably due to the difficulties of obtaining a sufficiently thin screen of high absorbing power.

The present theory can be applied to screens of any reflection coefficients. Hence, it could be tested without making a *perfectly absorbing* screen. The theory can also be applied to metal screens whose absorption cannot be neglected because of poor conductivity.

In addition it seems possible that the scattering by absorbing solid bodies whose dimensions are large compared to free space wave length can be treated in a similar way by first solving the problem for the perfect electrical and perfect magnetic conductor and finally adding both solutions together. One simple case where all the required calculations for the perfect conductor have been done is the scattering of a plane wave by a circular cylinder.

Acknowledgment

Essential parts of this paper have been achieved at the Eaton Electronics Research Laboratory of McGill University, Montreal. The author is indebted to its Director, Prof. G. A. Woonton, and to Dr. R. C. Spencer who supported the work by a grant of the U. S. Air Force, Cambridge Research Center, Cambridge, Mass

Side-Lobe Suppression by Pattern Multiplication*

RAYMOND JUSTICE†

***Summary*—It is shown that the properties of the radiation patterns of broadside and end-fire line radiators can be used to minimize the side-lobe level of the radiation patterns of uniformly excited rectangular arrays.**

I. Introduction

UNDER CERTAIN conditions the radiation pattern of a rectangular end-fire array can be represented as the product of the pattern of an end-fire line source and the pattern of a broadside line source. This well-known fact has some interesting consequences which may not be generally recognized.[1] If an end-fire line source with phase velocity equal to the velocity of light is uniformly energized, the highest side lobe in the resulting pattern is about 20 per cent of the maximum. Similarly, a uniformly energized broadside line source has a radiation pattern with the same side-lobe level. It follows that if the dimensions of a rectangular array are chosen so that these two patterns have the same beam width between first nulls, the product pattern will have side lobes that are not greater than 4 per cent of the maximum. The purpose of this paper is to illustrate the relationships between the length and width of such a rectangular array in order to best utilize the side-lobe suppressing characteristics of this multiplication process.

II. The End-Fire and Broadside Patterns

If the rectangle of the dimensions and orientation shown in Fig. 1 is covered by a continuous distribution of isotropic point sources that are uniformly energized in amplitude, in phase in the Y-direction and with a phase velocity in the Z-direction not greater that c, the speed of light, then the normalized radiation pattern in the Y-Z plane is given by[2]

* Manuscript received by the PGAP, June 26, 1955.

† Dept. to Electrical Engineering, Ohio State University, Columbus, Ohio.

[1] Final Engrg. Rep. 400-11, December 15, 1951, Antenna Lab., Ohio State Univ. Res. Found.; prepared under Contract AF 33(038)-9236, Air Res. and Dev. Command, Wright Air Dev. Center, Wright-Patterson AF Base, Ohio.

[2] Cf., for example, J. D. Kraus, "Antennas," McGraw-Hill Book Co., New York, N. Y., 1950.

$$|E(\theta)| = \frac{\sin X_1}{X_1}\frac{\sin X_2}{X_2}\left[\frac{X_2}{\sin X_2}\right]_{\theta=0} \quad (1)$$

where

$\frac{\sin X_1}{X_1}$ is the pattern of a broadside line source of length W'

$\frac{\sin X_2}{X_2}$ is the pattern of an end-fire line source of length L'

$$X_1 = \pi W \sin\theta$$

$$X_2 = \pi L(c/v - \cos\theta)$$

$$W = \frac{W}{\lambda}$$

$$L = \frac{L}{\lambda}$$

λ = free space wavelength

$$\frac{c}{v} = \frac{\text{velocity of light in vacuum}}{\text{phase velocity in } Z\text{-direction}}.$$

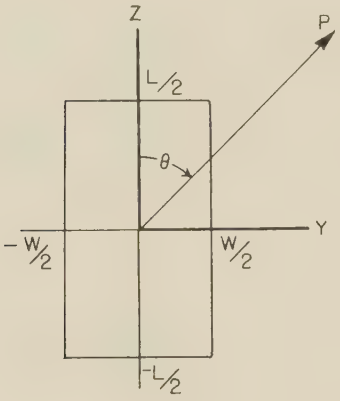


Fig. 1—Dimensions and orientation of array.

Curves of the individual patterns are given in Fig. 2 and Fig. 3 for a pair of values of W and L. Fig. 4 compares the end-fire patterns obtained from a given length for two values of c/v. Certain observations about these types of pattern can be made immediately on the basis of these examples. As mentioned in the introduction, the side lobes in the broadside pattern are at most 20 per cent of the maximum. For the end-fire pattern the same result holds if $c/v = 1$, but the side lobes increase rapidly in relative magnitude as the phase velocity is decreased. From the expression for the end-fire pattern and the fact that the function $\sin X/X$ has its largest secondary maxima at $X = 3\pi/2$, it follows that the side-lobe level of the end-fire pattern for values of $c/v \geqq 1$ is given by

Side-lobe level

$$= \left[\frac{\sin X_2}{X_2}\right]_{\theta=\theta m}\left[\frac{X_2}{\sin X_2}\right]_{\theta=0} 100 \text{ per cent of main beam} \quad (2)$$

where

$$\theta_m = \cos^{-1}\left[c/v - \frac{3}{2L}\right].$$

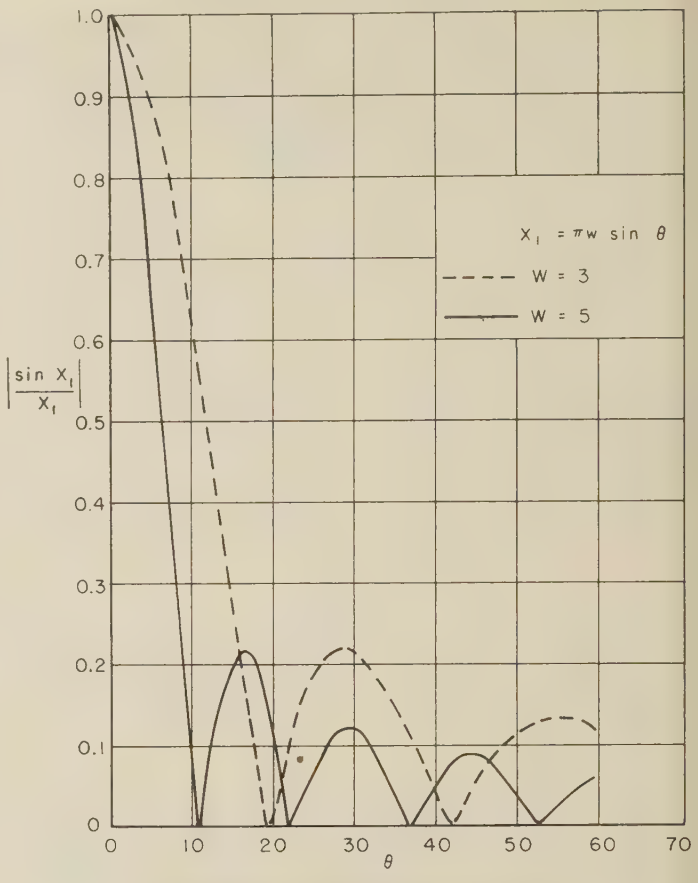


Fig. 2—Broadside patterns for $W = 3$ and $W = 5$.

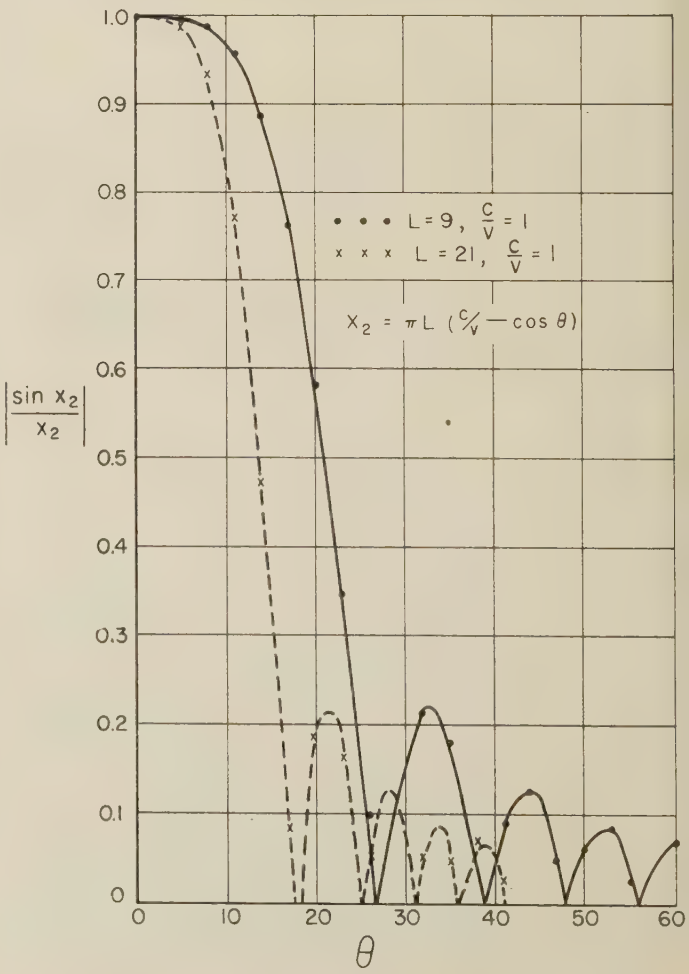


Fig. 3—End-fire patterns $L = 9$, $L = 19$, $c/v = 1$.

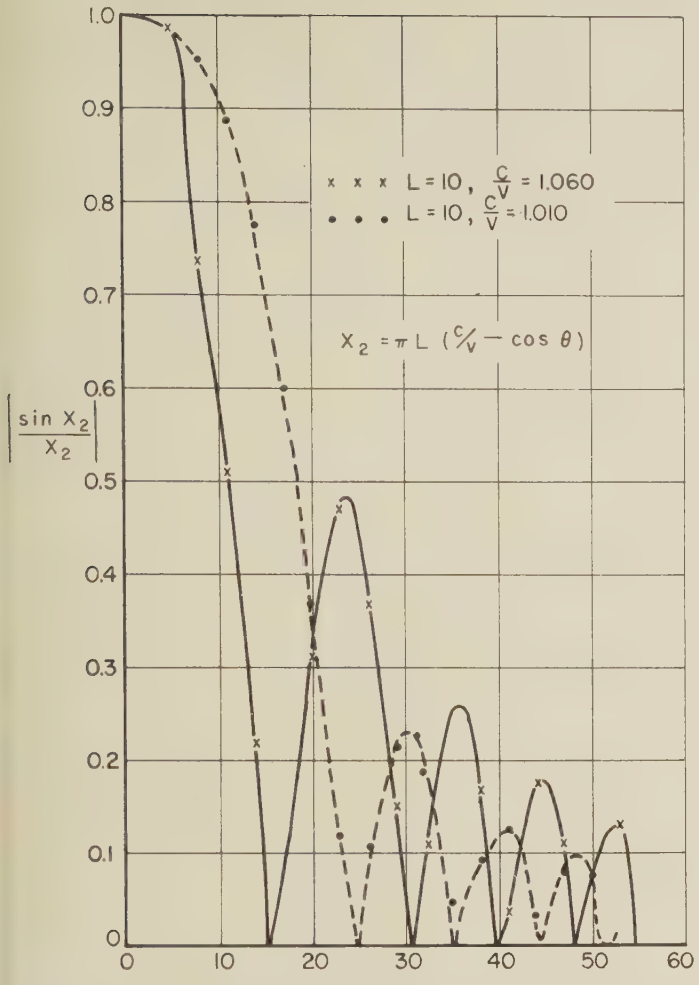


Fig. 4—Effect of c/v on end-fire pattern.

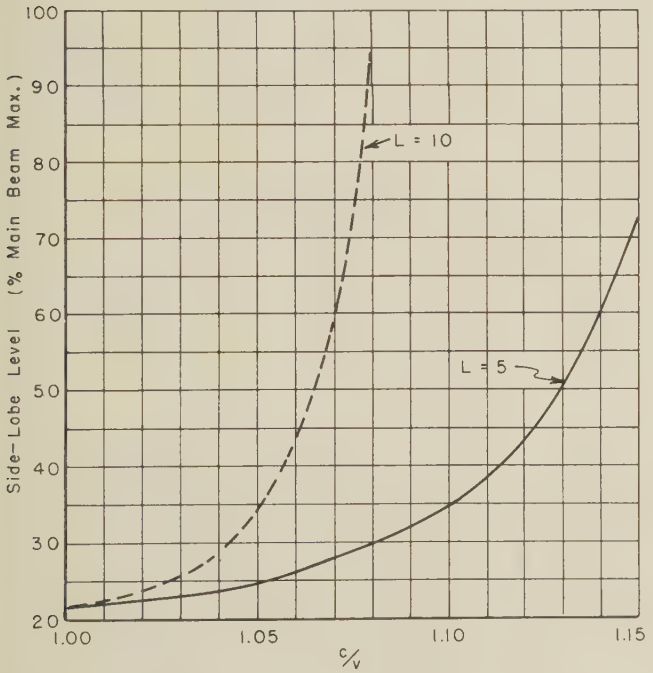


Fig. 5—Side-lobe level of end-fire pattern as a function of c/v.

Examples of this side-lobe level as a function of c/v for fixed L are given in Fig. 5. The rapid increase in side-lobe level for increasing c/v will ultimately limit the advantages of using such values of c/v for beam sharpening purposes (see Section IV).

The first zero of $\sin X/X$ occurs at $X=\pi$; hence the broadside and end-fire patterns have their first zeros at $\theta_{1W}=\sin^{-1}(1/W)$ and $\theta_{1L}=\cos^{-1}(c/v-1/L)$, respectively. This latter result illustrates the beam sharpening effect of increasing c/v, and it might be anticipated that a judicious use of this effect, while recalling the results of the preceding paragraph, could result in a narrower main beam for a given side-lobe level than could be achieved with $c/v=1$. Curves of θ_{1W} as a function of W and θ_{1L} as a function of L and c/v are given in Fig. 6. Observe that for L in excess of about 10, θ_{1L} does not decrease rapidly with increasing L. For later use a curve of $\theta_m=\cos^{-1}[1-3/2L]$ is included. θ_m is the location of the maximum of the first side lobe of the end-fire pattern for $c/v=1$.

III. The Case for $c/v=1$

The discussion in this section will be restricted to consideration of the patterns in the plane of the array, the Y-Z plane, when the phase velocity in the Z-direction is equal to c and the only parameters are the array dimensions, L and W. For this case both the end-fire and broadside patterns have side-lobe levels approximately equal to -13.2 db (20 per cent of main beam maximum). The condition of coincidence of first nulls, $\theta_{1W}=\theta_{1L}$, assures a pattern of the array with a side-lobe level of at most 4 per cent of the main beam. One might expect that proper choice of dimensions would give an even lower side-lobe level. That is, if the length of the array is fixed, the value of W for which minimum side lobes are achieved would be related to the fixed lengths. Analytical attempts to determine this relation have not been successful. However, by calculating large numbers of patterns and drawing graphs of side-lobe level as a function of one dimension for various values of the other dimension, it was possible to obtain sufficient data to estimate this relation, at least over the range of values for which patterns were calculated.

The accuracy of the result is limited by several factors. Since it was necessary to calculate large numbers of patterns it was not deemed practical to plot each pattern. The number chosen for the side-lobe level corresponds to the highest calculated value following the first minimum. The patterns were calculated at two-degree intervals of θ and the side-lobe levels determined from these calculations may contain considerable deviations from the true values. Furthermore, for a given value of one dimension, only a few values, up to seven, of the other dimensions were chosen for purposes of calculating the array pattern. As can be seen from the examples given in Fig. 7, it was frequently quite difficult with the available data to estimate with accuracy the relation between array dimensions for a minimum side-lobe level. However, considering all these things, the best estimate of the relation between L and W for minimum side lobes appears to be

$$\sin^{-1}\frac{1}{W}=\cos^{-1}\left(1-\frac{3}{2L}\right) \qquad (3)$$

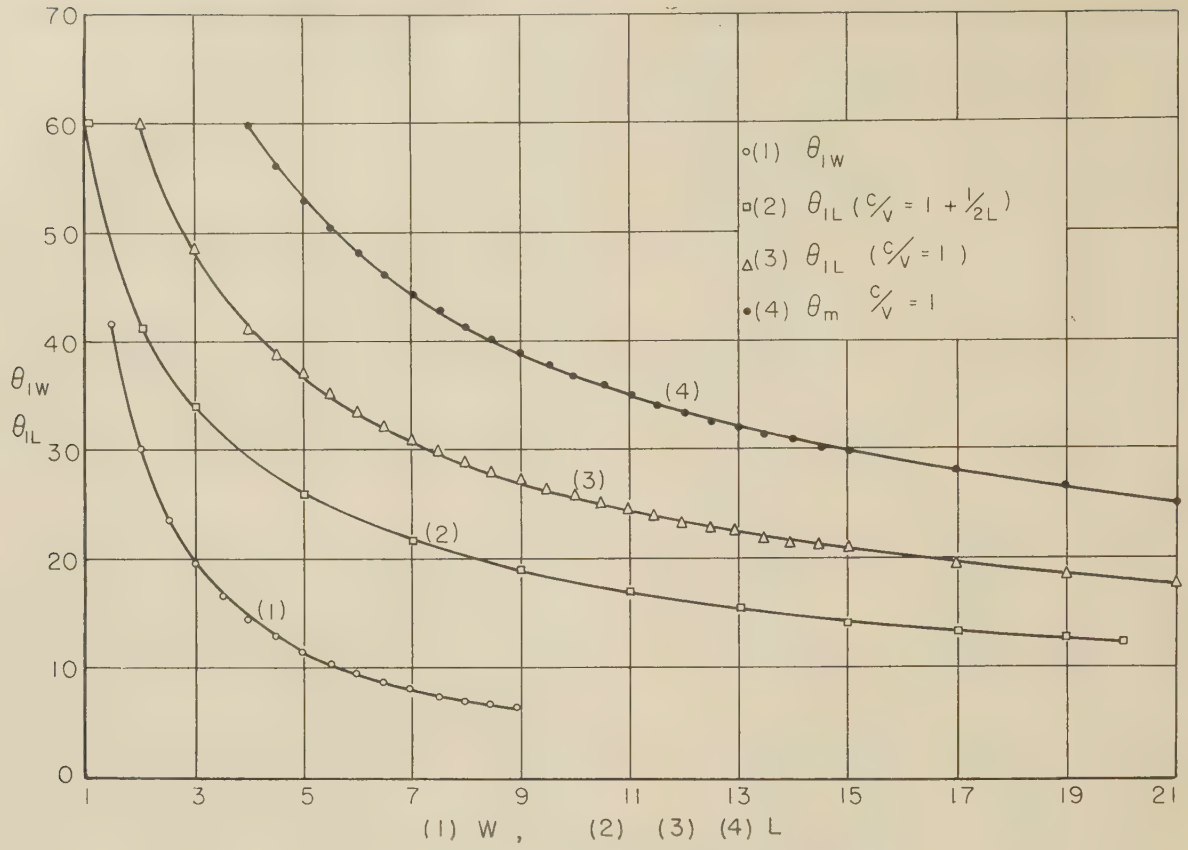


Fig. 6—Characteristics of broadside and end-fire patterns.

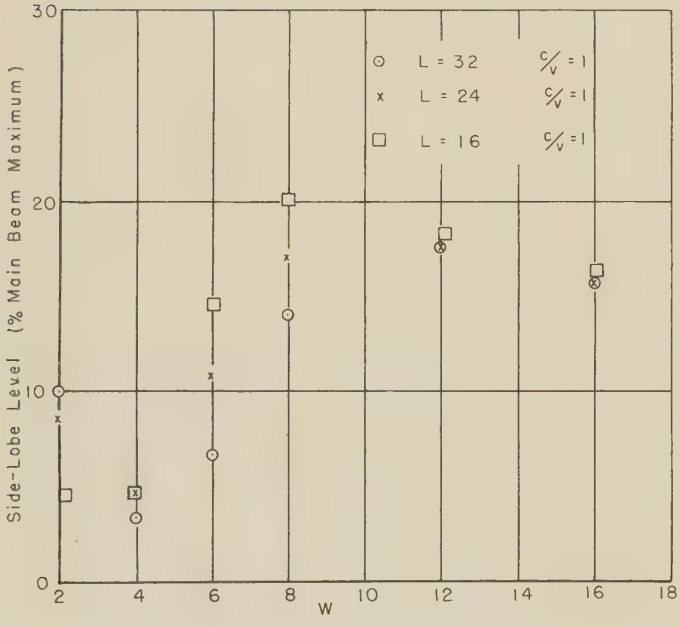


Fig. 7—Side-lobe levels as a function of W.

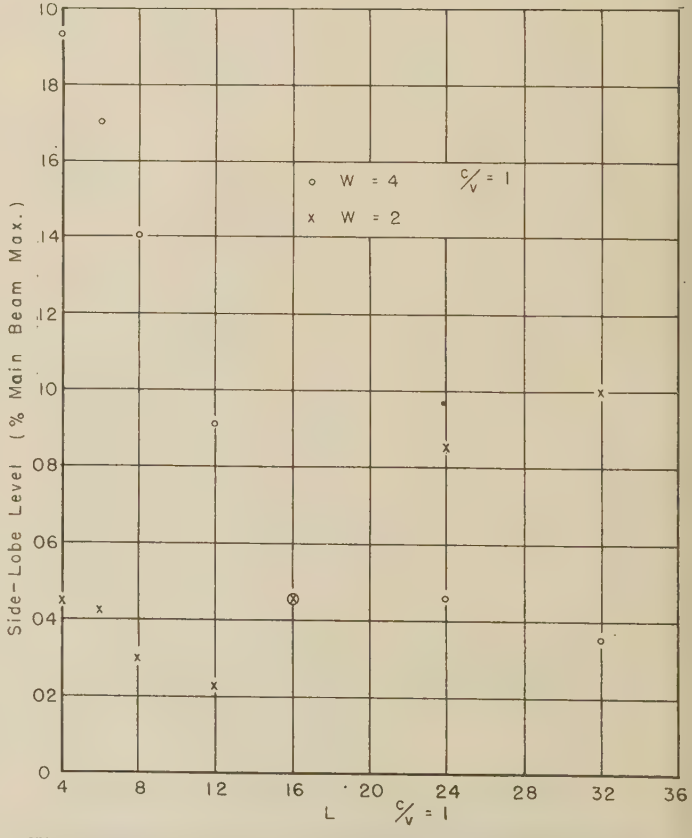


Fig. 8—Effect of W on rate of change of side-lobe level with L.

for the case $c/v=1$. This relation corresponds to the first null of the broadside pattern occurring at the same angle at which the end-fire pattern has its first (and highest) secondary maximum. In general it may be qualitatively stated that for this case ($c/v=1$) the side-

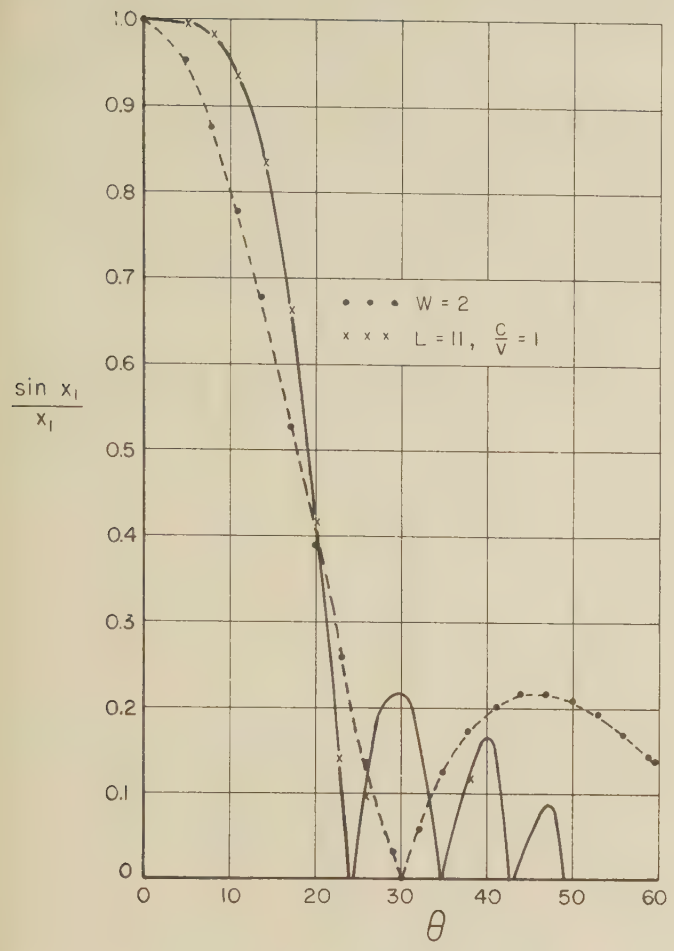


Fig. 9—End-fire and broadside patterns.

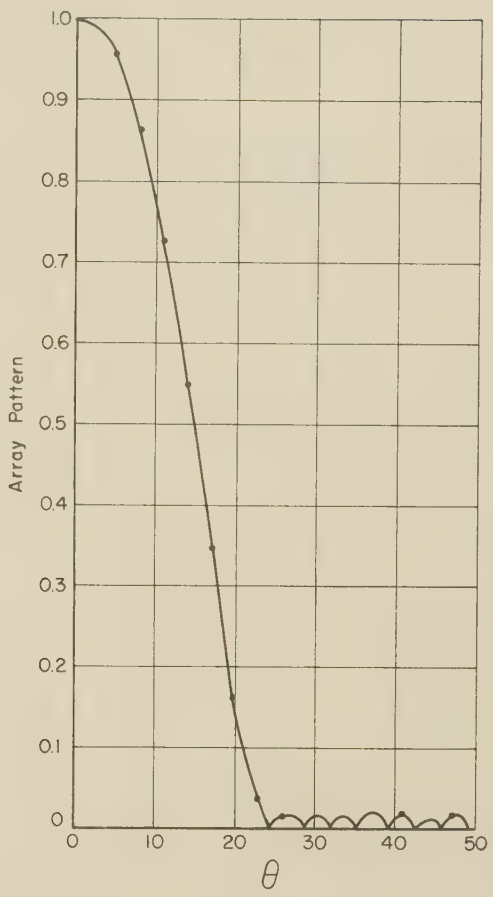


Fig. 10—Pattern of an array with dimensions satisfying (3).

lobe level does not vary rapidly with L in the neighborhood of the value which satisfies (3). The variation with L is slower for the larger values of W as exemplified by comparing the curves of Fig. 8 for $W=2$ and $W=4$. This is to be expected since the value of L which satisfied (3) for the larger values of W is in the region where the curve of θ_{1L} as a function of length is varying quite slowly. An example of the results obtainable by choosing L and W to satisfy (3) is given in Figs. 9 and 10 where the two factors of the array pattern and the array pattern itself are given.

IV. The Case for $c/v \geqq 1$

In this case the array dimensions are fixed and the variable parameter is the phase velocity v. If this velocity is decreased to a value less than c, the end-fire pattern has a narrower main beam but higher side lobes. The problem is to determine when the disadvantages of the latter effect essentially outweigh the former.

The analytical difficulties encountered in the case for $c/v=1$ are compounded by the introduction of phase velocity as an additional variable. In this case it was necessary to compute even larger numbers of patterns than previously in order to obtain sufficient data to indicate what relations may exist between the three parameters for a minimum side-lobe level. Since no definitive result was obtained, a qualitative discussion must suffice.

For the narrower widths, say $W<2$, it generally turned out that the optimum phase velocity for a given length L in the sense of minimum side lobes was about the same as that required by the Hansen-Woodyard condition for maximum directivity of an end-fire line source of length L.[3] In the present notation this phase velocity satisfies

$$c/v = 1 + \frac{1}{2L}. \tag{4}$$

For wider arrays the broadside pattern may have sufficient influence on the end-fire pattern to cause the optimum c/v to be greater than that given by (4).

The limitations on the accuracy of the determination of side-lobe level as a function of one of the three parameters are the same here as in the case for $c/v=1$. Samples of side-lobe level as a function of c/v for a few dimensions are given in Figs. 11 and 12.

[3] W. W. Hansen and J. R. Woodward, "A new principle in directional antenna design," Proc. IRE, vol. 26, pp. 333–345; March, 1938.

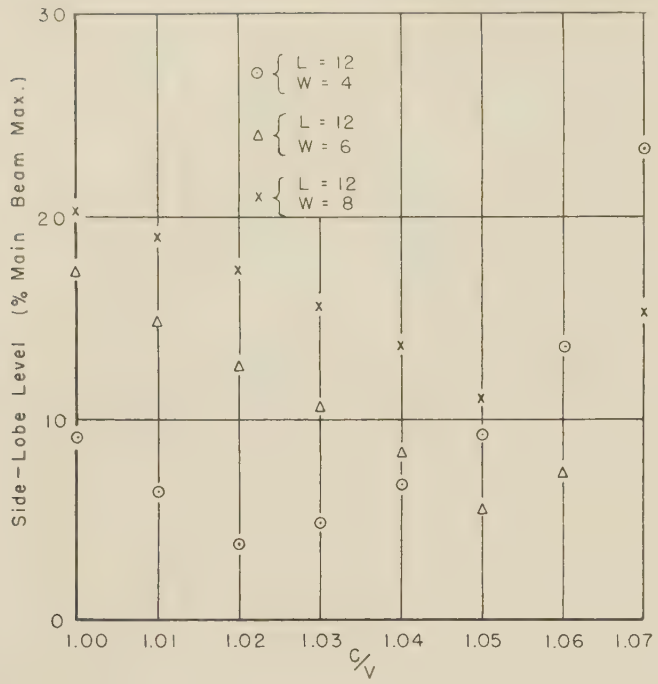


Fig. 11—Side-lobe levels as functions of c/v.

V. Conclusions

It is not intended to imply that the optimum combinations of L, W and c/v will give a particularly low side-lobe level in all cases. The conclusion to be drawn here is that for $c/v=1$ and one dimension fixed, there is an optimum value of the other dimension. For this case the side lobes under optimum conditions are not more than 4 per cent of the main beam maximum and are usually appreciably smaller than this. If the array dimensions are fixed, then there is an optimum value of c/v which will give minimum side lobes for an array of these dimensions. In this case it is not possible to give an exact upper bound to the side-lobe levels, but on the basis of calculations made to date it may be asserted that optimum conditions will assure side lobes less than 10 per cent of the main beam maximum.

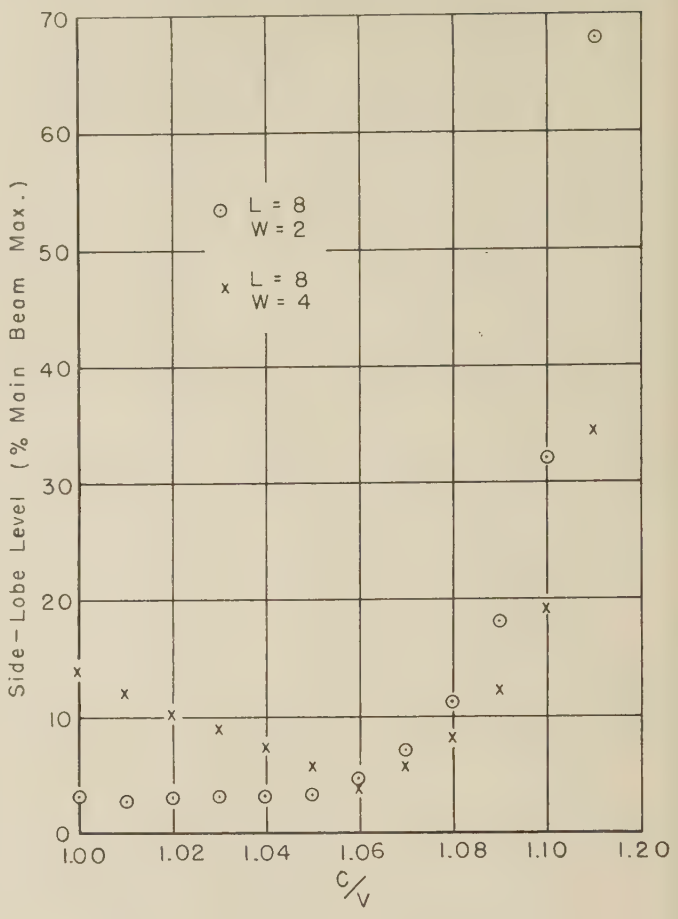


Fig. 12—Side-lobe levels as functions of c/v.

Acknowledgment

The author gratefully acknowledges the contributions of C. H. Walter and V. H. Rumsey. Mrs. F. Phillips performed the computations. The work described in this paper was carried out under Contract AF 18(600) 85 (E.O. Nos. 112-110 SR-6F2, 112-105 SA-9B) Air Research and Development Command, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio.

On the Conductance of Slots*

JAMES R. WAIT†

Summary—**The external radiation conductance G of narrow half-wave axial and transverse slots on a circular cylinder, a ribbon, a knife edge, and a right-angle wedge are computed. The method is based on finding the total power radiated assuming a sinusoidal variation of voltage along the slot. The theoretical values of G for the ribbon of width $2d$ are applied to the case of a half-wave slot cut in the narrow face of a rectangular waveguide which has a flush mounted plate or flange of width $2d$.**

* Manuscript received by the PGAP, August 25, 1955.
† National Bureau of Standards, Boulder, Colo.

IN THE DESIGN of slot antennas it is usually assumed for purposes of calculating the conductance that the surface on which the slot is cut is flat and infinite in extent. In practical radiating systems the surface is often curved to an appreciable extent such as when the slot is located in the leading edge of an aircraft wing. In other situations, the slot may be cut in the broad or narrow faces of a rectangular waveguide which have dimensions comparable with the wavelength.

It is the purpose of this paper to show how the conductance is affected by the curvature and the finite extent of the surface on which it is cut. To facilitate the solution of the problem it is necessary to idealize the shape of the surface which is assumed to be perfectly conducting. The simplest case is probably the circular cylinder of infinite length of radius a.

The conductance G is now obtained by calculating the total power P radiated from the slot and then dividing by $V^2/2$ where V is the voltage at the center of the slot. Furthermore, it is assumed that the voltage distribution along the slot is sinusoidal. Choosing a spherical coordinate system (R, θ, ϕ) such that $\theta = 0, \pi$ corresponds to the axis of the cylinder, the far zone fields are given by [1, 2, 3]

$$E_\theta = \frac{e^{-ikR}}{R} V P_\theta(\theta, \phi) \tag{1}$$

and

$$E_\phi = \frac{e^{-ikR}}{R} V P_\phi(\theta, \phi) \tag{2}$$

where $k = 2\pi/\text{wavelength}$ and P_θ and P_ϕ are pattern factors expressible in terms of a Fourier Series as follows:

$$P_\theta(\theta, \phi) = \sum_{m=0}^{\infty} A_m \cos m\phi + a_m \sin m\phi \tag{3}$$

and

$$P_\phi(\theta, \phi) = \sum_{m=0}^{\infty} B_m \sin m\phi + b_m \cos m\phi \tag{4}$$

where A_m, a_m, B_m and b_m are coefficients which are a function of θ. For a narrow half-wave axial slot centered at $\phi = 0$, $\theta = \pi/2$ and located on the surface of the cylinder [2],

$$A_m = a_m = B_m = 0 \tag{5}$$

and

$$b_m = \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\pi^2 ka \sin^2\theta} \frac{\epsilon_m i^m}{H_m^{(2)\prime}(ka \sin\theta)} \tag{6}$$

where

$$\epsilon_0 = 1, \qquad \epsilon_m = 2 (m \neq 0)$$

and where $H_m^{(2)}$ is the Hankel function of the second kind and the prime indicates a derivative with respect to the argument. In the case of the narrow half-wave transverse slot centered at the same point [3],

$$A_m = \frac{ka}{\pi^2 \sin\theta} \frac{\epsilon_m i^{m-1} \cos(m\pi/2ka)}{H_m^{(2)\prime}(ka \sin\theta)\left[(ka)^2 - m^2\right]}, \tag{7}$$

$$B_m = \frac{ka \cot\theta}{\pi^2 \sin\theta} \cdot \frac{2m i^{m-1} \cos(m\pi/2ka)}{H_m^{(2)\prime}(ka \sin\theta)\left[(ka)^2 - m^2\right]}, \tag{8}$$

and $a_m = b_m = 0$. The conductance is now obtained from

$$G = \frac{2P}{V^2} = \frac{1}{\eta_0}\int_0^\pi d\theta \int_0^{2\pi} d\phi \left[E_\theta E_\theta^* + E_\phi E_\phi^*\right] R^2 \sin\theta \qquad R \to \infty \tag{9}$$

where $\eta_0 = 120\pi$ and the asterisk indicates a complex conjugate. The integration with respect to ϕ can be carried out to yield

$$G = \frac{\pi}{\eta_0}\int_0^\pi \sum_{m=0}^{\infty} \left[|A_m|^2 (2/\epsilon_m) + |a_m|^2 + |B_m|^2 + |b_m|^2 (2/\epsilon_m)\right] \sin\theta d\theta. \tag{10}$$

In the case of the axial half-wave slot

$$G = \frac{2}{\eta_0 \pi^3 k^2 a^2}\int_0^\pi \frac{\cos^2\left[(\pi/2)\cos\theta\right]}{\sin^3\theta} \sum_{m=0}^{\infty} \frac{\epsilon_m}{|H_m^{(2)\prime}|^2} d\theta \tag{11}$$

and in the case of the transverse half-wave slot

$$G = \frac{2}{\eta_0 \pi^3}\int_0^\pi \frac{1}{\sin\theta} \sum_{m=0}^{\infty} \frac{\epsilon_m \cos^2(m\pi/2ka)}{\left[(ka)^2 - m^2\right]^2} \cdot \left[\frac{(ka)^2}{|H_m^{(2)}|^2} + \frac{m^2 \cot\theta}{|H_m^{(2)\prime}|^2}\right] d\theta \tag{12}$$

where the argument of the Hankel functions is understood to be $ka \sin\theta$. Because of the presence of the square of the moduli of the Hankel functions in the denominators, the infinite series converge quite rapidly. The terms in the summation with order m greater than about ka are negligible. The integration with respect to θ cannot apparently be carried out in closed form. The integrand, however, turns out to be slowly varying function of θ and the conductance G can therefore be obtained readily by a graphical integration. A slight difficulty occurs in (12) for the case of the transverse slot since the integrand has a logarithmic singularity at $\theta = 0$ and π. The contribution to G in this case in the regions $\theta = 0$ to δ and $\theta = \pi - \delta$ to π is denoted by G_δ and for δ small is given by

$$G_\delta \simeq \frac{1}{30\pi^4}\int_0^\delta \frac{d\theta}{\sin\theta (ka)^2 |H_0^{(2)}(ka \sin\theta)|^2} \tag{13}$$

where only the term for $m = 0$ is significant if $ka \sin\delta \simeq ka\delta \ll 1$. Using the small argument approximation for Hankel functions, it follows that

$$G_\delta \simeq \frac{1}{30\pi^4 (ka)^2}\int_0^\delta \frac{d\theta}{\left\{1 + (4/\pi^2)\left[\log(\Gamma ka\theta/2)\right]^2\right\}\theta} \tag{14}$$

with $\log \Gamma = 0.5772$. The latter integral can be evaluated in closed form to yield

$$G_\delta \simeq \frac{1}{60\pi^3 (ka)^2}\left[\tan^{-1}\left[\frac{\log\frac{\Gamma ka\delta}{2}}{(\pi/2)}\right] + \pi/2\right] \tag{15}$$

where the inverse tangent is a negative angle less than $\pi/2$. This contribution is then added to the portion of the integral obtained by the graphical procedure to obtain the conductance of the transverse slot.

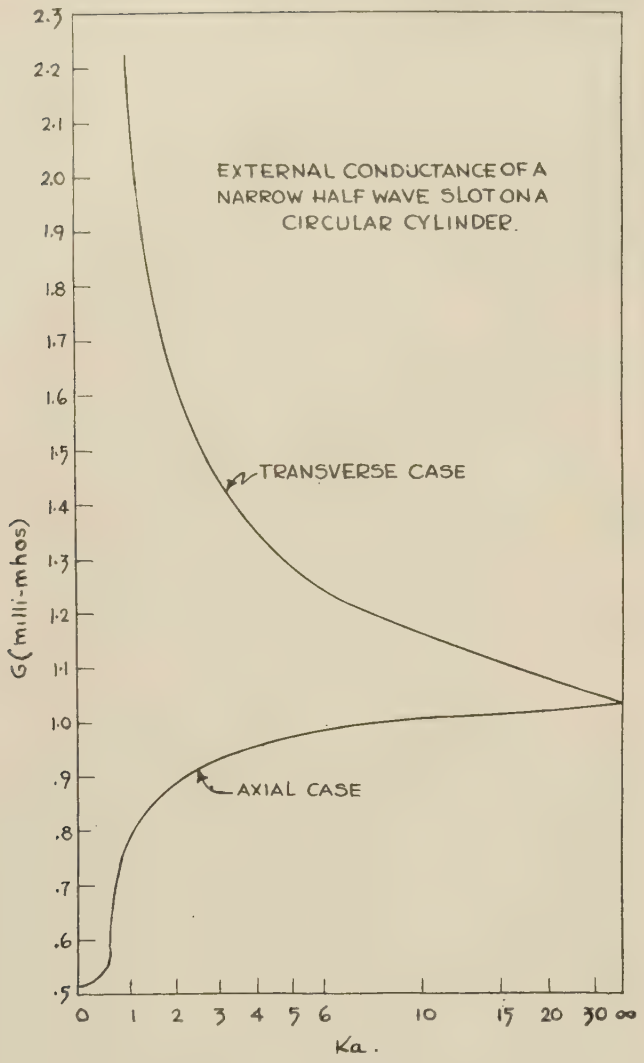


Fig. 1—The computed external radiation conductance for axial and transverse half-wave slots on a circular cylinder as a function of the circumference of the cylinder in wavelengths.

The conductance for both axial and transverse half-wave slots on a circular cylinder is shown plotted in Fig. 1 with an exponential scale for the abscissa. It can be seen that the effect of curvature is very pronounced. In the case of the axial slot the conductance varies between the limits 0.515 and 1.03 millimhos from small to large cylinders respectively. It can be readily deduced from Babinet's principle that the 1.03 corresponds to the conductance of a thin half-wave slot cut in a flat surface of infinite extent and perfect conductivity. In the case of the transverse slot the conductance increases as ka becomes smaller.

The above method can also be employed to calculate the conductance of slots on elliptic cylinders. Essentially, the technique is to integrate the total power radiated from a prescribed slot on a perfectly conducting cylinder of elliptic cross-section. The functions involved in this case are angular and radial Mathieu functions [4] which play the role of the exponential and Bessel functions respectively in the circular cylinder case. Some calculations for short slots on elliptic cylinders for special values of the ellipticity have been published by Wong [5]. Further results for axial half-wave slots were given recently by Wait and Walpole [6] for the case when the slot was located at the center of one side of ribbon or elliptic cylinder of vanishing minor axis. In this latter paper results were also given for a half-wave slot located on one side of a half-plane and parallel to the edge. The result for the half-plane or knife edge can be easily extended to the case when the slot is on the surface of a wedge [7]. For example, if the half-wave slot is parallel to and distance ρ from the apex of a right-angled wedge, the conductance G is given by

$$G = \frac{80}{\eta_0^2}\int_0^{\pi} \frac{\cos^2\left[(\pi/2)\cos\theta\right]}{\sin\theta} \cdot \sum_{m=0,1,2\cdots}^{\infty} \epsilon_m \left[J_{2m/3}(k\rho\sin\theta)\right]^2 d\theta \qquad (16)$$

where $J_{2m/3}$ is the Bessel function of the first type of order $2m/3$.

Using the above-mentioned formulas, the conductance G for narrow half-wave axial slots is shown plotted in Fig. 2 for the cases where the slot is on a knife edge, right angled wedge, thin ribbon and circular cylinder.

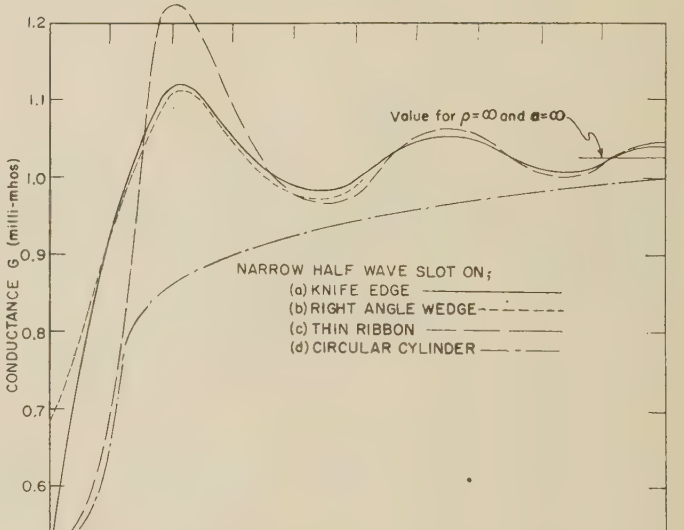


Fig. 2—The computed external radiation conductance of a half-wave axial slot on a knife edge, right angled wedge, thin ribbon and circular cylinder.

In the case of the ribbon the distance ρ corresponds to the half-width of the ribbon. The results for the axial slot on the circular cylinder has been replotted in Fig. 2 for purposes of comparison.

In Fig. 3 the conductance G is plotted for the case where the half-wave slot is moved from one side of the

ribbon to the other for a fixed width ($kd=8$). The corresponding result for the slot on the knife edge is also shown for comparison. The close agreement between the results for the knife edge and the ribbon when the slot is near the edge is reassuring. Of course as the slot approaches the middle of the ribbon, the other edge influences the result.

To check the results for the ribbon experimentally, the slot was cut in the narrow face of a rectangular X-band waveguide. Then thin rectangular metal plates of various widths $2d$ were mounted centrally and flush with the narrow face of the guide. The equivalent shunt conductance g of the slot normalized to the characteristic admittance of the guide can then be written in the following form

$$g(kd) = \frac{480}{73\pi}\frac{c}{b}\frac{\lambda_g}{\lambda}\cos^2\left[\frac{\pi\lambda}{2\lambda_g}\right]\frac{1.03}{G} \qquad (17)$$

where b and c are the inner narrow and broad dimensions of the guide respectively, λ_g and λ are the guide and free space wavelengths, respectively, and G is the external or aperture conductance in millimhos of the thin half-wave slot. This formula is obtained by a straightforward extension of Stevenson's analysis [7] who assumed the external surface of the waveguide to be infinite in extent. In other words, when kd approaches infinity, G approaches 1.03 millimhos, and the formula for g is identical to Stevenson's result.

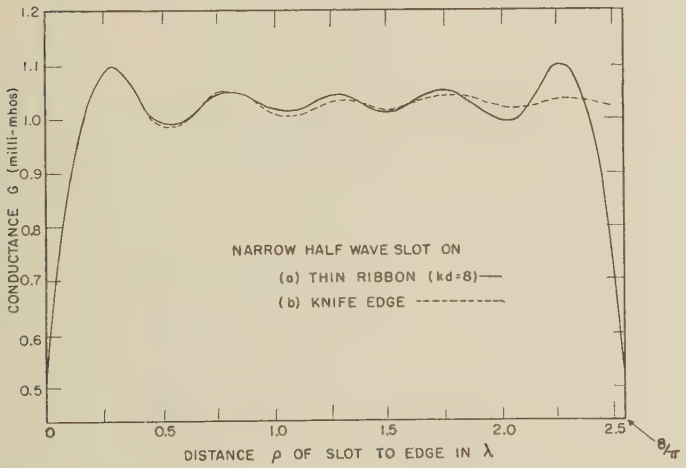


Fig. 3—The computed external radiation conductance of a half-wave axial slot as a function of distance to the edge of a ribbon of fixed width with the result for the knife edge shown for comparison.

The experimental values of $g(kd)$ are shown plotted in Fig. 4 expressed as an increment to the conductance $g(\infty)$ for a very large sheet. The details of the technique of measurement are standard, and more detailed results can be found in a published report [8]. The theoretical results shown are based on the calculations for the ribbon which of course neglect the presence of the waveguide behind the plate. For larger plates the agreement is good, and for small plate widths the inadequacy of the theory is evident.

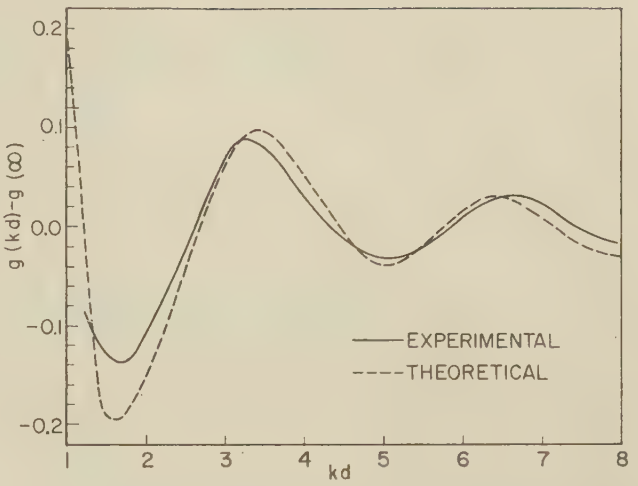


Fig. 4—The theoretical and experimental curves for the normalized shunt conductance of the half-wave longitudinal slot in the narrow face of an X-band rectangular waveguide with a flange of width $2d$.

The above computed data for slot conductances should prove useful in the design of slotted antenna arrays on cylindrical structures. For surfaces of more complex shape the results show at least the error involved in assuming the surface to be flat and infinite in extent.

Acknowledgment

I would like to thank W. A. Pope who assisted with the calculations and H. V. Cottony who read the manuscript.

Bibliography

[1] Silver, S., and Saunders, W. K., "The External Field Produced by a Slot in an Infinite Circular Cylinder." *Journal of Applied Physics*, Vol. 21 (February, 1950), pp. 153–158.

[2] Wait, J. R., "Radiation Characteristics of Axial Slots on a Circular Conducting Cylinder." *Wireless Engineer*, Vol. 32 (December, 1955), pp. 316–322.

[3] Wait, J. R., and Kahana, S., "Calculated Patterns of Circumferential Slots on a Circular Conducting Cylinder." *Canadian Journal of Technology*, Vol. 33 (January, 1955), pp. 77–97.

[4] Wait, J. R., "Field Produced by an Arbitrary Slot on an Elliptic Cylinder." *Journal of Applied Physics*, Vol. 26 (April, 1955), pp. 458–463.

[5] Wong, J. Y., "Radiation Conductance of Axial and Transverse Slots in Cylinders of Elliptic Cross Section." Proceedings of the IRE, Vol. 41 (September, 1953), pp. 1172–1176.

[6] Wait, J. R., and Walpole, R. E., "Calculated Radiation Characteristics of Slots Cut in Metal Sheets." *Canadian Journal of Technology*, Vol. 33 (May, 1955), pp. 211–227.

[7] Wait, J. R., and Kahana, S., "Radiation from a Slot on a Cylindrically Tipped Wedge." *Canadian Journal of Physics*, Vol. 34 (November, 1954), pp. 714–721.

[8] Watson, W. H., "The Physical Principles of Waveguide Transmission and Antenna Systems." Oxford University Press (1947). (A. F. Stevenson's theory of slots in rectangular waveguides is excellently summarized in this reference.)

[9] Frood, D. G., and Wait, J. R., "An Investigation of Slot Radiators Cut in Metal Plates." *Proceedings of The Institute of Electrical Engineers*, Vol. 103 (January, 1956), pp. 103–110.

A Method of Analyzing Coupled Antennas of Unequal Sizes*

C. A. LEVIS† AND C. T. TAI†

Summary—**The impedance parameters of coupled antennas are calculated by a variational method. The unknown coefficients pertaining to the trial functions for the current distributions on various elements are shown to satisfy a set of linear equations involving the excitation conditions, but the final formulas for the impedance parameters involve only the trial functions and the geometry. The method is directly applicable to the analysis of Yagi-Uda antennas, and the technique can be extended to coupled slots inside a waveguide or coupled radiating slots excited by waveguides.**

Introduction

SIGNIFICANT contributions to the problem of coupled cylindrical antennas of unequal sizes have been made by Uda and Mushiake[1] in their study of Yagi-Uda antennas by an extension of Hallén's method for a single antenna.[2] A similar method was used by King and Harrison[3] in their treatment of two coupled antennas of identical size. The method described in this paper is an extension of Storer's variational method, originally used for the treatment of a single cylindrical antenna.[4,5] An extension proposed by Begovich[6] gives variational formulas for coupled slots, provided they are small electrically. Although no numerical values are yet available for the solution of the Yagi-Uda antenna by this new method, we should like to present here a general derivation of the essential formulas involved in the method.

Simultaneous Integral Equations and the Variational Expression

We will illustrate the method by considering two coupled antennas of unequal sizes. Corresponding formulas for n coupled antennas are given in the Appendix. Fig. 1 shows a sketch of the two coupled antennas. The approximate integral equations governing the current distributions on the two antennas are given by

$$V_1\,\delta(z-0) = \int_{-l_1}^{l_1} I_1(z')G_{11}(z-z')dz' + \int_{-l_2}^{l_2} I_2(z')G_{12}(z-z')dz' \qquad (1)$$

$$V_2\,\delta(z-0) = \int_{-l_1}^{l_1} I_1(z')G_{12}(z-z')dz' + \int_{-l_2}^{l_2} I_2(z')G_{22}(z-z')dz', \qquad (2)$$

where

V_1, V_2 = voltages applied to the two antennas,

$I_1(z)$, $I_2(z)$ = current distributions on the two antennas to be determined,

$\delta(z-0)$ = Dirac delta function defined at $z=0$,

$$G_{11}(z-z') = \frac{jk\eta}{4\pi}\left(1+\frac{1}{k^2}\frac{\partial^2}{\partial z^2}\right)\frac{e^{-jk\sqrt{(z-z')^2+a_1^2}}}{\sqrt{(z-z')^2+a_1^2}},$$

$$\eta = \left(\frac{\mu}{\epsilon}\right)^{1/2} = 120\pi \text{ ohms},$$

$k = 2\pi/\lambda$,

a_1 = radius of the first antenna.

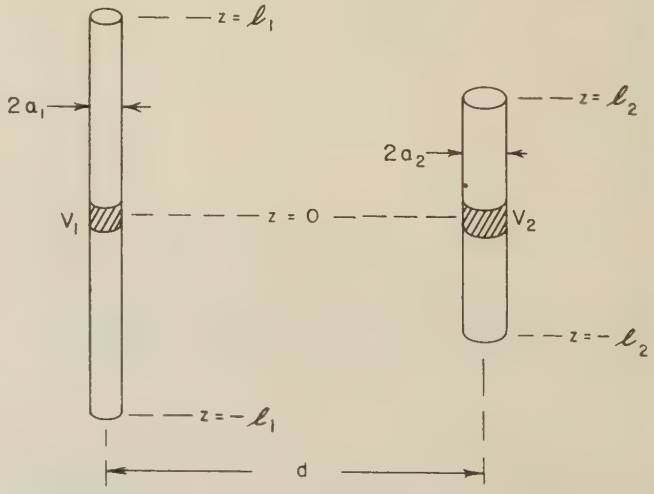


Fig. 1—Two coupled antennas of unequal sizes.

The functions $G_{22}(z-z')$ and $G_{12}(z-z')$ are obtained by replacing a_1 in $G_{11}(z-z')$ by a_2 and d respectively, where a_2 denotes the radius of the second antenna and d the distance between them.

* Manuscript received by the PGAP, August 10, 1955. The research reported in this paper was sponsored in part by the Air Res. and Dev. Command, Wright Air Dev. Ctr., Dayton, Ohio.

† Antenna Lab., Dept. of Elect. Eng., Ohio State Univ., Columbus, Ohio.

[1] S. Uda and Y. Mushiake, "Yagi-Uda Antenna," Sasaki Publishing Co., Sendai, Japan; 1954.

[2] E. Hallén, "Theoretical investigations into the transmitting and receiving qualities of antennae," *Nova Acta*, Uppsala, ser. IV, vol. 11, November 11, 1938.

[3] R. King and C. W. Harrison, "Mutual and self-impedances for coupled antennas," *J. Appl. Phys.*, vol. 15, pp. 481–495; June, 1944.

[4] J. E. Storer, "Variational solution to the problem of the symmetrical cylindrical antenna," Tech. Report no. 101, Cruft Laboratory, Harvard Univ., Cambridge, Mass.; 1950.

[5] C. T. Tai, "A variational solution to the problem of cylindrical antennas," Tech. Report no. 12, Aircraft Radiation Systems Laboratory, Stanford Research Institute, Stanford, Calif.; 1950.

[6] N. A. Begovich, "Waveguide excited slot radiators," Tech. Memo no. 219, Hughes Aircraft Co. Res. and Dev. Labs., 1950.

From the linearity of Maxwell's equations, it can be shown that the currents and voltages at the antenna terminals satisfy the relationships

$$V_1 = Z_{11}I_1(0) + Z_{12}I_2(0), \tag{3}$$

$$V_2 = Z_{12}I_1(0) + Z_{22}I_2(0). \tag{4}$$

The parameters Z_{11}, Z_{12}, Z_{22} characterize the impedance behavior of the antennas and are to be calculated. Eqs. (1) and (2) may therefore be rewritten as

$$[Z_{11}I_1(0) + Z_{12}I_2(0)]\delta(z - 0) = \int_{-l_1}^{l_1} I_1(z')G_{11}(z - z')dz' + \int_{-l_2}^{l_2} I_2(z')G_{12}(z - z')dz', \tag{5}$$

$$[Z_{12}I_1(0) + Z_{22}I_2(0)]\delta(z - 0) = \int_{-l_1}^{l_1} I_1(z')G_{12}(z - z')dz' + \int_{-l_2}^{l_2} I_2(z')G_{22}(z - z')dz'. \tag{6}$$

If we multiply (5) by $I_1(z)$ and integrate from $-l_1$ to l_1, multiply (6) by $I_2(z)$ and integrate from $-l_2$ to l_2, and add the results, we obtain

$$Z_{11}I_1^2(0) + 2Z_{12}I_1(0)\,I_2(0) + Z_{22}I_2^2(0) = \iint I_1G_{11}I_1 + 2\iint I_1G_{12}I_2 + \iint I_2G_{22}I_2 \tag{7}$$

where the specific arguments of the integrals at the right have been omitted for convenience; *i.e.*,

$$\iint I_mG_{mn}I_n = \int_{-l_m}^{l_m}\int_{-l_n}^{l_n} I_m(z)G_{mn}(z - z')I_n(z')dz'dz. \tag{8}$$

Eq. (7) is the basis of a variational solution for the parameters Z_{11}, Z_{12}, Z_{22}. It may be used in the following way. We choose the currents corresponding to three arbitrary but nonidentical excitation conditions (denoted by single, double, and triple dots, respectively) and substitute them in (7). The result is three simultaneous equations, viz.,

$$Z_{11}\dot{I}_1^2(0) + 2Z_{12}\dot{I}_1(0)\dot{I}_2(0) + Z_{22}\dot{I}_2^2(0) = \iint \dot{I}_1G_{11}\dot{I}_1 + 2\iint \dot{I}_1G_{12}\dot{I}_2 + \iint \dot{I}_2G_{22}\dot{I}_2, \tag{9}$$

$$Z_{11}\ddot{I}_1^2(0) + 2Z_{12}\ddot{I}_1(0)\ddot{I}_2(0) + Z_{22}\ddot{I}_2^2(0) = \iint \ddot{I}_1G_{11}\ddot{I}_1 + 2\iint \ddot{I}_1G_{12}\ddot{I}_2 + \iint \ddot{I}_2G_{22}\ddot{I}_2, \tag{10}$$

$$Z_{11}\dddot{I}_1^2(0) + 2Z_{12}\dddot{I}_1(0)\dddot{I}_2(0) + Z_{22}\dddot{I}_2^2(0) = \iint \dddot{I}_1G_{11}\dddot{I}_1 + 2\iint \dddot{I}_1G_{12}\dddot{I}_2 + \iint \dddot{I}_2G_{22}\dddot{I}_2. \tag{11}$$

This system can be solved for the parameters Z_{11}, Z_{12}, Z_{22} provided the determinant of their coefficients does not vanish, and this is precisely what was meant above by "nonidentical" excitation conditions. (It can be shown that it is necessary and sufficient to require that $I_1(0)/I_2(0)$ have a distinct value for each of the three excitation conditions. In practice it may be convenient, when two such distributions are known, to obtain the third by superposition.) The solutions of (9), (10), (11) will be the true values of Z_{11}, Z_{12}, Z_{22} provided the currents are true currents; *i.e.*, currents which satisfy (1) and (2). The solutions have the further property of being stationary with respect to variations of the currents about such true distributions. This property is demonstrated below.

To demonstrate the variational nature of the simultaneous solutions of (9), (10), (11), we take variations of the currents about their true values. If we make use of the symmetric nature of the G_{mn} we get from (9)

$$\delta Z_{11}\dot{I}_1^2(0) + 2\delta Z_{12}\dot{I}_1(0)\dot{I}_2(0) + \delta Z_{22}\dot{I}_2^2(0) + 2\delta\dot{I}_1(0)[Z_{11}\dot{I}_1(0) + Z_{12}\dot{I}_2(0)] + 2\delta\dot{I}_2(0)[Z_{12}\dot{I}_1(0) + Z_{22}\dot{I}_2(0)] = 2\Big[\iint \delta\dot{I}_1G_{11}\dot{I}_1 + \iint \delta\dot{I}_1G_{12}\dot{I}_2 + \iint \delta\dot{I}_2G_{12}\dot{I}_1 + \iint \delta\dot{I}_2G_{22}\dot{I}_2\Big]. \tag{12}$$

Since $\dot{I}_1$ and $\dot{I}_2$ were to be varied about the true distributions, (5) and (6) must be satisfied by these distributions. Multiplying (5) by $\delta\dot{I}_1(z)$ and integrating from $-l_1$ to l_1 gives

$$\delta\dot{I}_1(0)[Z_{11}\dot{I}_1(0) + Z_{12}\dot{I}_2(0)] = \iint \delta\dot{I}_1G_{11}\dot{I}_1 + \iint \delta\dot{I}_1G_{12}\dot{I}_2. \tag{13}$$

A similar expression can be obtained by multiplying (6) by $\delta\dot{I}_2(z)$ and integrating from $-l_2$ to l_2. When these results are used in (12) we find

$$\delta Z_{11}\dot{I}_1^2(0) + 2\delta Z_{12}\dot{I}_1(0)\dot{I}_2(0) + \delta Z_{22}\dot{I}_2^2(0) = 0. \tag{14}$$

The same procedure, applied to (10) and (11), gives relations identical to (14) except that the single dot is replaced by a double or triple dot, respectively. These three equations may be solved simultaneously for δZ_{11}, δZ_{12}, δZ_{22}. Since the determinant of the coefficients is the same as that in (9) to (11), the only solution is

$$\delta Z_{11} = \delta Z_{12} = \delta Z_{22} = 0. \tag{15}$$

This guarantees that if approximate currents which differ from the true currents by errors of the first order are used in (9) to (11), the calculated parameters will differ from the true values only by errors of the second order.

Approximate Solution using Trial Functions of Two Terms

Unfortunately it is generally difficult to guess good approximations for the currents $\dot{I}_1(z)$, $\dot{I}_2(z)$, etc., directly. Instead it is customary to represent the currents as a sum of functions with coefficients to be determined by some auxiliary method. To discuss the final step of the formulation, which leads to formulas for numerical calculation, let us consider the approximate solution

which is based on the use of trial functions consisting of two terms. This is actually the case that we intend to apply to the analysis of Yagi-Uda antennas. We assume

$$I_1(z) = A_1 f_1(z) + A_2 f_2(z) \tag{16}$$

$$I_2(z) = A_3 f_3(z) + A_4 f_4(z), \tag{17}$$

where the $f_k(z)$ are four known functions which are believed to be a good representation of the current distributions on the two antennas. For example, we can choose $f_1(z)$ and $f_3(z)$ to be $\sin[k(l_1-|z|)]$ and $\sin[k(l_2-|z|)]$, which are the distribution functions for "infinitely thin" antennas. If we use these functions alone, the variational solution would yield the same result as that obtained by the induced emf method.[7] The terms $f_2(z)$ and $f_4(z)$ are therefore necessary in order to obtain an improvement over the old method. As in the case of the single antenna[4,5] we can use functions of the form $\cos(kl)-\cos(kz)$ or $k(l-|z|)\cos[k(l-|z|)]$ for $f_2(z)$ and $f_4(z)$. It should be noted that the coefficients A_k will be, in general, functions of the excitation. The currents corresponding to particular excitations will be characterized by particular values of the A_k; for instance we may write

$$\dot{I}_1(z) = \dot{A}_1 f_1(z) + \dot{A}_2 f_2(z), \tag{18}$$

and so forth.

When the currents are assumed to be of this form, the stationary expression, (7), becomes

$$Z_{11}[A_1a_1 + A_2a_2]^2 + 2Z_{12}[A_1a_1 + A_2a_2][A_3a_3 + A_4a_4] + Z_{22}[A_3a_3 + A_4a_4]^2 = \sum_i \sum_k A_i A_k \gamma_{ik}, \tag{19}$$

where

$$\gamma_{ik} = \int_{-l_m}^{l_m} \int_{-l_n}^{l_n} f_i(z) G_{mn}(z - z') f_k(z') dz' dz = \gamma_{ki} \tag{20}$$

with

$$m = \begin{cases} 1 \text{ for } i = 1, 2 \\ 2 \text{ for } i = 3, 4 \end{cases} \tag{21}$$

$$n = \begin{cases} 1 \text{ for } k = 1, 2 \\ 2 \text{ for } k = 3, 4 \end{cases}. \tag{22}$$

The constants a_1, a_2, a_3, a_4 are abbreviations for $f_1(0)$, $f_2(0)$, $f_3(0)$, and $f_4(0)$, respectively. Because of the stationary character of (7), as shown in (15), we shall determine the coefficients by differentiating (19) with respect to the A_k and requiring

$$\frac{\partial Z_{11}}{\partial A_k} = \frac{\partial Z_{12}}{\partial A_k} = \frac{\partial Z_{22}}{\partial A_k} = 0. \tag{23}$$

The result is a set of equations linear in the A_k;

$$a_1[Z_{11}(A_1a_1 + A_2a_2) + Z_{12}(A_3a_3 + A_4a_4)] = A_1\gamma_{11} + A_2\gamma_{12} + A_3\gamma_{13} + A_4\gamma_{14}, \tag{24}$$

$$a_2[Z_{11}(A_1a_1 + A_2a_2) + Z_{12}(A_3a_3 + A_4a_4)] = A_1\gamma_{12} + A_2\gamma_{22} + A_3\gamma_{23} + A_4\gamma_{24}, \tag{25}$$

$$a_3[Z_{12}(A_1a_1 + A_2a_2) + Z_{22}(A_3a_3 + A_4a_4)] = A_1\gamma_{13} + A_2\gamma_{23} + A_3\gamma_{33} + A_4\gamma_{34}, \tag{26}$$

$$a_4[Z_{12}(A_1a_1 + A_2a_2) + Z_{22}(A_3a_3 + A_4a_4)] = A_1\gamma_{14} + A_2\gamma_{24} + A_3\gamma_{34} + A_4\gamma_{44}. \tag{27}$$

It should be remembered that the variational expression, (7) or equivalently (19), was to be used to calculate Z_{11}, Z_{12}, Z_{22} by specifying three excitation conditions, thus obtaining three equations of the form of (19) with the A_k replaced by $\dot{A}_k$, $\ddot{A}_k$, and $\dddot{A}_k$. Thus the set of linear equations above is in reality a shorthand for three sets; *i.e.*, for solutions to (24) to (27) we seek three sets of constants $\dot{A}_k$, $\ddot{A}_k$, $\dddot{A}_k$, which together with a single set of Z_{11}, Z_{12}, Z_{22} will satisfy the equations. Any such solutions will also satisfy the variational expression (19), since this expression can be obtained from (24) to (27) by simple algebraic manipulations.

To solve for a particular set of the A_k, we must specify the excitation, and a convenient way of doing this is to set

$$Z_{11}(A_1a_1 + A_2a_2) + Z_{12}(A_3a_3 + A_4a_4) = V_1, \tag{28}$$

$$Z_{12}(A_1a_1 + A_2a_2) + Z_{22}(A_3a_3 + A_4a_4) = V_2, \tag{29}$$

and then to assign particular values $\dot{V}_1$, $\dot{V}_2$; $\ddot{V}_1$, $\ddot{V}_2$; $\dddot{V}_1$, $\dddot{V}_2$ to the constants V_1 and V_2. When (28) and (29) are substituted into (24) to (27), we obtain the following solutions for the A_k:

$$A_k = [V_1(a_1\Delta_{1k} + a_2\Delta_{2k}) + V_2(a_3\Delta_{3k} + a_4\Delta_{4k}]/\Delta, \tag{30}$$

where Δ denotes the determinant $|\gamma_{ik}|$ and Δ_{ik} the cofactor of the element γ_{ik}. Eq. (30) again represents three sets of equations, the $\dot{A}_k$, for example, being given by it with particular values $\dot{V}_1$, $\dot{V}_2$ assigned to V_1 and V_2.

The final step is to substitute these values of A_k into (28) and (29). From (28) we obtain, after collecting terms and noting that $\Delta_{ik}=\Delta_{ki}$ [because of (20)],

$$V_1[Z_{11}(a_1^2\Delta_{11} + 2a_1a_2\Delta_{12} + a_2^2\Delta_{22}) + Z_{12}(a_1a_3\Delta_{13} + a_2a_3\Delta_{23} + a_1a_4\Delta_{14} + a_2a_4\Delta_{24}) - \Delta] + V_2[Z_{11}(a_1a_3\Delta_{13} + a_1a_4\Delta_{14} + a_2a_3\Delta_{23} + a_2a_4\Delta_{24}) + Z_{12}(a_3^2\Delta_{33} + 2a_3a_4\Delta_{34} + a_4^2\Delta_{44})] = 0. \tag{31}$$

This equation can again be interpreted as a set which must hold when the particular values $\dot{V}_1$, $\dot{V}_2$ etc. are substituted; this can be the case only if the coefficients of V_1 and V_2 vanish individually. Setting these coefficients equal to zero, we obtain two equations in the unknowns Z_{11}, Z_{12} which have the solutions

$$Z_{11} = \Delta(a_3^2\Delta_{33} + 2a_3a_4\Delta_{34} + a_4^2\Delta_{44})/M, \tag{32}$$

$$Z_{12} = -\Delta(a_1a_3\Delta_{13} + a_1a_4\Delta_{14} + a_2a_3\Delta_{23} + a_2a_4\Delta_{24})/M, \tag{33}$$

[7] P. S. Carter, "Circuit relations in radiating systems and applications to antenna problems," PROC. IRE, vol. 20, p. 1006; June, 1932.

where

$$M = \begin{vmatrix} a_1^2\Delta_{11} + 2a_1a_2\Delta_{12} + a_2^2\Delta_{22} & a_1a_3\Delta_{13} + a_1a_4\Delta_{14} + a_2a_3\Delta_{23} + a_2a_4\Delta_{24} \\ a_1a_3\Delta_{13} + a_1a_4\Delta_{14} + a_2a_3\Delta_{23} + a_2a_4\Delta_{24} & a_3^2\Delta_{33} + 2a_3a_4\Delta_{34} + a_4^2\Delta_{44} \end{vmatrix}. \quad (34)$$

Similar manipulation on (29) yields the same value for Z_{12}, and in addition

$$Z_{22} = \Delta(a_1^2\Delta_{11} + 2a_1a_2\Delta_{12} + a_2^2\Delta_{22})/M. \quad (35)$$

These are the formulas from which calculations for the Yagi-Uda antenna are to be made.

It should be noted that instead of the circuit relations of (3), (4) we could have written

$$I_1(0) = V_1Y_{11} + V_2Y_{12} \quad (36)$$

$$I_2(0) = V_1Y_{12} + V_2Y_{22}. \quad (37)$$

Parameters Y_{11}, Y_{12}, Y_{22} are related to impedances by

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12} & Z_{22} \end{bmatrix}^{-1} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{12} & Y_{22} \end{bmatrix}. \quad (38)$$

In our case, the Y's have a simpler form than the Z's, *viz.*,

$$Y_{11} = (a_1^2\Delta_{11} + 2a_1a_2\Delta_{12} + a_2^2\Delta_{22})/\Delta \quad (39)$$

$$Y_{12} = (a_1a_3\Delta_{13} + a_1a_4\Delta_{14} + a_2a_3\Delta_{23} + a_2a_4\Delta_{24})/\Delta \quad (40)$$

$$Y_{22} = (a_3^2\Delta_{33} + 2a_3a_4\Delta_{34} + a_4^2\Delta_{44})/\Delta. \quad (41)$$

Finally we may mention that the linear set of equations for the A_k, and the final results (39) to (41), can be obtained by applying V. H. Rumsey's Reaction Principle[8,9] to our problem. This method indeed leads to the admittance result more directly, but was not chosen here because the concepts are not as widely known.

Appendix: Extension to n Antennas

The extension to n antennas with an arbitrary number of trial functions for each antenna is straightforward. The geometry of two typical elements of the array is shown in Fig. 2. Instead of going through the proof in detail, the corresponding steps are only indicated here, but the numbering is kept the same; *e.g.*, (7a) here corresponds to (7) in the two-antenna argument. The definition of G_{rs} is obtained by replacing a_1 in G_{11} with a_r for $r=s$, and with d_{rs} for $r\neq s$. In analogy to the two-antenna case, we now have

$$V_k\,\delta(z-0) = \sum_{r=1}^{n}\int_{-l_r}^{l_r} I_r(z')G_{kr}(z-z')dz' \quad (1a)$$

$$V_k = \sum_{r=1}^{n} Z_{kr}I_r(0) \quad (3a)$$

$$\sum_{r=1}^{n} Z_{kr}I_r(0)\delta(z-0) = \sum_{r=1}^{n}\int_{-l_r}^{l} I_r(z')G_{kr}(z-z')dz' \quad (5a)$$

$$\sum_{r=1}^{n}\sum_{s=1}^{n} Z_{sr}I_s(0)I_r(0) = \sum_{r=1}^{n}\sum_{s=1}^{n}\int_{-l_r}^{l_r}\int_{-l_s}^{l_s} I_r(z')G_{sr}(z-z')I_s(z)dz'dz \quad (7a)$$

subject to $Z_{ts}=Z_{st}$. In analogy with (9) to (11), we here have to write $\frac{1}{2}n(n+1)$ linearly independent equations.

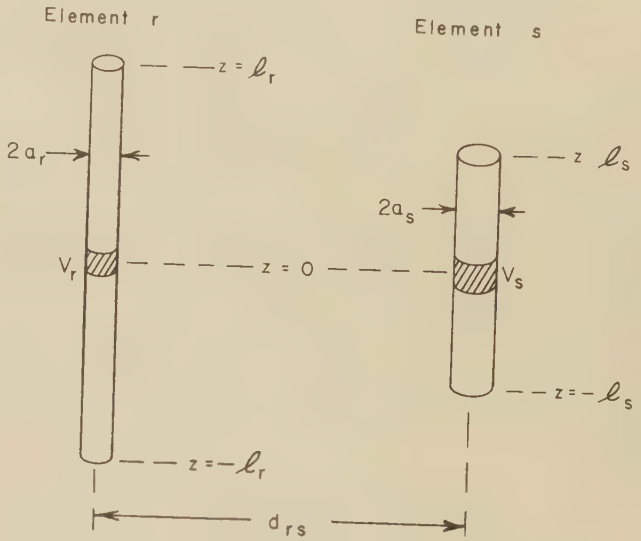


Fig. 2—Typical elements of the n-antenna array.

If the Z_{sr} are obtained as solutions of these equations, it follows that

$$\delta Z_{sr} = 0, \qquad s, r = 1, 2, \cdots, n. \quad (15a)$$

Suppose we want to represent each current I_k as a sum of trial functions f_q, with coefficients A_q to be determined by some auxiliary method. We write

$$I_k(z) = \sum_{q=m_k}^{n_k} A_q f_q(z) \quad (16a)$$

with $m_1=1$ and $m_{k+1}=n_k+1$. This numbers the functions consecutively: The functions $f_1, f_2, \cdots, f_{n_1}$ contribute to I_1, the functions $f_{n_1+1}, f_{n_1+2}, \cdots, f_{n_2}$ contribute to I_2, etc. We also abbreviate by writing a_k for $f_k(0)$. From (7a) we now obtain

$$\sum_{r=1}^{n}\sum_{s=1}^{n} Z_{sr}\sum_{q=m_r}^{n_r} A_qa_q\sum_{p=m_s}^{n_s} A_pa_p = \sum_{q=1}^{n_n}\sum_{p=1}^{n_n} A_qA_p\gamma_{qp}, \quad (19a)$$

where

$$\gamma_{qp} = \int_{-l_q}^{l_q}\int_{-l_p}^{l_p} f_q(z)G_{uv}(z-z')f_p(z')dz'dz = \gamma_{pq} \quad (20a)$$

with u and v chosen to that $m_u\leqq q\leqq n_u$, $m_v\leqq p\leqq n_v$.

Again we obtain the coefficients by differentiating with respect to the A_k and requiring $\partial Z_{sr}/\partial A_k=0$ for all s, r, and k. Noting the symmetrical properties of the Z_{sr} and γ_{qp}, we obtain

[8] V. H. Rumsey, "The reaction concept in electromagnetic theory," *Phys. Rev.*, vol. 94, pp. 1483–1491; June, 1954.

[9] C. A. Levis and R. E. Webster, "Mutual Admittances and Annular Slots," Project Report 486-19, Antenna Laboratory, The Ohio State University Research Foundation; prepared under Contract AF 18(600)85, Air. Res. and Dev. Command, Wright Air Dev. Ctr., Wright-Patterson Air Force Base, Ohio; 1954.

$$a_k \sum_{r=1}^{n} Z_{sr} \sum_{q=m_r}^{n_r} A_q a_q = \sum_{q=1}^{n_n} A_q \gamma_{qk}, \quad k = 1, 2, \cdots, n_n \quad (24a)$$

with s determined by $m_s \leqq k \leqq n_s$. To specify excitation, it is convenient to let

$$\sum_{r=1}^{n} Z_{sr} \sum_{q=m_r}^{n_r} A_q a_q = V_s \qquad s = 1, 2, \cdots, n \qquad (28a)$$

and then to assign specific values to the V_s. Use of (28a) in (24a) gives

$$a_k V_s = \sum_{q=1}^{n_n} A_q \gamma_{qk} \qquad \begin{matrix} k = 1, 2, \cdots, n_n \\ m_s \leqq k \leqq n_s. \end{matrix}$$

This system can be solved for the A_q

$$A_q = \frac{1}{\Delta} \sum_{s=1}^{n} V_s \sum_{k=m_s}^{n_s} a_k \Delta_{qk}, \qquad (30a)$$

where Δ is the determinant $|\gamma_{qk}|$ (assumed nonvanishing) and Δ_{qk} is the cofactor of the element γ_{qk}. It is now necessary to insure that these values also satisfy (28a). Substitution in this equation gives

$$\sum_{t=1}^{n} V_t \left(\sum_{r=1}^{n} Z_{sr} \alpha_{tr} - \Delta \delta_{ts} \right) = 0 \quad s = 1, \cdots, n, \qquad (31a)$$

where

$$\alpha_{tr} = \sum_{q=m_r}^{n_r} \sum_{k=m_t}^{n_t} a_q a_k \Delta_{qk}$$

and

$$\delta_{ts} = \begin{cases} 1 \text{ for } t = s \\ 0 \text{ for } t \neq s. \end{cases}$$

Eq. (31a) must hold for $\frac{1}{2}n(n+1)$ sets of nonidentical excitations. This is possible only if the coefficients of the V_t vanish; *i.e.*,

$$\sum_{r=1}^{n} Z_{sr} \alpha_{tr} = \Delta \delta_{ts} \qquad \begin{matrix} t = 1, 2, \cdots, n \\ s = 1, 2, \cdots, n. \end{matrix}$$

For fixed s, we can consider this a set of n equations in the n unknowns Z_{sr} having the solutions

$$Z_{sr} = \frac{\Delta}{|\alpha_{sr}|} A_{sr} = Z_{rs}, \qquad (32a)$$

where $|\alpha_{sr}|$ denotes the determinant of the α_{sr} (assumed nonzero) and A_{sr} the cofactor of the element α_{sr} in the determinant. Again it turns out that an admittance representation yields a simpler formula; *viz.*,

$$Y_{sr} = \frac{\alpha_{sr}}{\Delta} = \left(\sum_{q=m_r}^{n_r} \sum_{k=m_s}^{n_s} a_q a_k \Delta_{qk} \right) \Big/ \Delta \qquad \begin{matrix} s = 1, \cdots, n \\ r = 1, \cdots, n. \end{matrix} \qquad (39a)$$

Radiation Characteristics of the Spherical Luneberg Lens*

EDWARD H. BRAUN†

LUNEBERG [1] has shown that a spherically symmetric region in which the dielectric constant varies as $2-r^2$ (where $r \leqq 1$ is the normalized radius) will focus rays from a point source at $r=1$ into a plane phase front having a propagation vector pointing along the diameter passing through the source point.

Construction of such a dielectric region has been impractical in the past because of lack of availability of suitable dielectric materials. Recently, however, materials have been developed which make it feasible to build such a lens. This so-called "Spherical Luneberg Lens" has a number of practical applications, based principally on its high degree of symmetry. These applications as well as the details of construction however will not be described in this paper, the principal purpose of which is to discuss the far field radiation characteristics of the lens.

Initially we will assume only that the amplitudes of the electric and magnetic fields are arbitrary known functions of position over the surface of the lens, while the phases are such as to result in a plane phase front perpendicular to the axis of the lens when the rays are extended out through the aperture.

The pattern is given by the expression

$$\vec{E}(R, \theta, \phi) = -j \frac{\omega \mu}{4\pi} \frac{e^{-jkR}}{R} \vec{F}(\theta, \phi)$$

where

$$\vec{F}(\theta, \phi) = \iint \left[\hat{\rho} \times \vec{H} - (\hat{\rho} \times \vec{H}) \cdot \hat{R}\hat{R} - \sqrt{\frac{\epsilon}{\mu}} (\hat{\rho} \times \vec{E}) \times \hat{R} \right] e^{jk\vec{\rho} \cdot \hat{R}} dS.$$

* Manuscript received by PGAP, May 27, 1955; revised manuscript received November 29, 1955.

† Microwave Antennas and Components Branch, Electronics Division, Naval Res. Lab., Washington, D. C.

Here $\vec{\rho}$ is a vector from the center of the lens to a point on the surface, ρ is a unit vector in this direction, $\hat{R}$ is a unit vector in the direction of the field point (See Fig. 1). Taking the dot product with $\hat{\theta}$ and $\hat{\Phi}$ (unit vectors at the field point) we obtain,

$$F_\theta = \int\int \left[(\hat{\rho} \times \vec{H}) \cdot \hat{\theta} - \sqrt{\frac{\epsilon}{\mu}} (\hat{\rho} \times \vec{E}) \cdot \hat{\Phi} \right] e^{ik\vec{\rho}\cdot\hat{R}} dS \quad (1)$$

$$F_\Phi = \int\int \left[(\hat{\rho} \times \vec{H}) \cdot \hat{\Phi} + \sqrt{\frac{\epsilon}{\mu}} (\hat{\rho} \times \vec{E}) \cdot \hat{\theta} \right] e^{ik\vec{\rho}\cdot\hat{R}} dS. \quad (2)$$

We take the electric field at any point on the lens to be linearly polarized in the x direction and given by $\vec{E} = E(\vartheta', \phi')\hat{x}$, where ϑ' and ϕ' are polar coordinates on the surface of the lens.

Hence

$$E_{\vartheta'} = E\hat{x} \cdot \hat{\vartheta}' = E \cos\vartheta' \cos\phi'$$
$$E_{\phi'} = E\hat{x} \cdot \hat{\phi}' = -E \sin\phi'.$$

The magnetic field at any point on the lens is given by $\vec{H} = H(\vartheta', \phi')\hat{y}$, hence

$$H_{\vartheta'} = H\hat{y} \cdot \hat{\vartheta}' = H \cos\vartheta' \sin\phi'$$

and

$$H_{\phi'} = H\hat{y} \cdot \hat{\phi}' = H \cos\phi'$$

$$\hat{\rho} \times \vec{H} \cdot \hat{\theta} = H \cos\vartheta' \sin\phi' \hat{\phi}' \cdot \hat{\theta} - H \cos\phi' \hat{\vartheta}' \cdot \hat{\theta}.$$

Now:

$$\hat{\phi}' \cdot \hat{\theta} = \cos\theta \sin(\Phi - \phi')$$
$$\hat{\vartheta}' \cdot \hat{\theta} = \sin\vartheta' \sin\theta + \cos\vartheta' \cos\theta \cos(\Phi - \phi')$$
$$\hat{\rho} \times \vec{E} \cdot \hat{\Phi} = E \cos\vartheta' \cos\phi' \hat{\phi}' \cdot \hat{\Phi} + E \sin\phi' \hat{\vartheta}' \cdot \hat{\Phi}$$
$$\hat{\phi}' \cdot \hat{\Phi} = \cos(\Phi - \phi')$$
$$\hat{\vartheta}' \cdot \hat{\Phi} = \cos\vartheta' \sin(\phi' - \Phi)$$
$$\hat{\rho} \times \vec{H} \cdot \hat{\Phi} = H \cos\vartheta' \sin\phi' \cos(\Phi - \phi') - H \cos\vartheta' \cos\phi' \sin(\phi' - \Phi)$$
$$(\hat{\rho} \times \vec{E}) \cdot \hat{\theta} = E \cos\vartheta' \cos\phi' \cos\theta \sin(\Phi - \phi') + E \sin\phi' [\sin\vartheta' \sin\theta + \cos\vartheta' \cos\theta \cos(\Phi - \phi')]$$
$$\vec{\rho} \cdot \hat{R} = a \cos\vartheta' \cos\theta + a \sin\vartheta' \sin\theta \cos(\Phi - \phi').$$

The distance to a plane parallel to the (x, y) plane and passing through the point $(0, 0, a)$ is $d = a - \vec{\rho} \cdot \hat{z} = a - a\cos\vartheta'$.

Hence we may write E and H as

$$E = E_0(\vartheta', \phi') e^{jka} e^{-jka\cos\vartheta'}$$
$$H = H_0(\vartheta', \phi') e^{jka} e^{-jka\cos\vartheta'}$$

where the phases of $E_0(\vartheta', \phi')$ and $H_0(\vartheta', \phi')$ are constant over the sphere.

Substituting into (1), and using the identity:

$$\cos\phi' \cos(\Phi - \phi') - \sin\phi' \sin(\Phi - \phi') = \cos\Phi$$

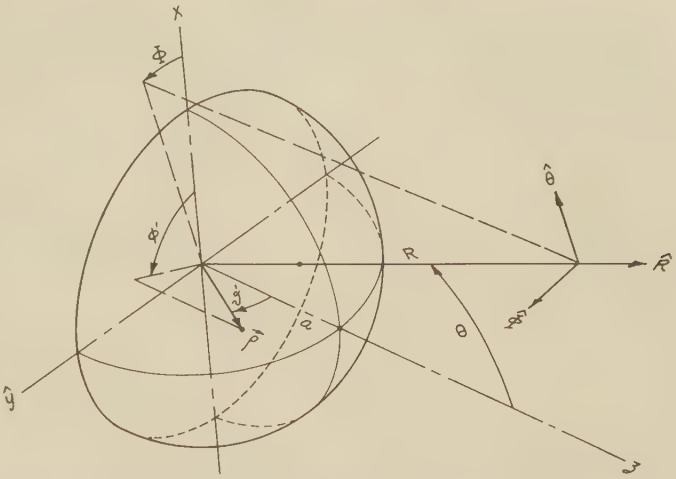


Fig. 1

we obtain:

$$F_\theta = -e^{jka} a^2 \int_0^{\pi/2} \int_0^{2\pi} \left[\left(H_0 \cos\theta + \sqrt{\frac{\epsilon}{\mu}} E_0 \right) \cdot \cos\vartheta' \cos\Phi + H_0 \sin\theta \sin\vartheta' \cos\phi' \right] \cdot e^{-jka[2\sin^2\theta/2 \cos\vartheta' - \sin\theta \sin\vartheta' \cos(\Phi - \phi')]} \cdot \sin\vartheta' d\vartheta' d\phi'. \quad (3)$$

Using (2) and the identity

$$\sin\phi' \cos(\Phi - \phi') + \cos\phi' \sin(\Phi - \phi') = \sin\Phi$$

we obtain:

$$F_\Phi = a^2 e^{jka} \int_0^{\pi/2} \int_0^{2\pi} \left\{ \left[H_0 + \sqrt{\frac{\epsilon}{\mu}} E_0 \cos\theta \right] \cdot \cos\vartheta' \sin\Phi + \sqrt{\frac{\epsilon}{\mu}} E_0 \sin\theta \sin\phi' \sin\vartheta' \right\} \cdot e^{-jka[2\sin^2\theta/2 \cos\vartheta' - \sin\theta \sin\vartheta' \cos(\Phi - \phi')]} \cdot \sin\vartheta' d\vartheta' \phi'. \quad (4)$$

Letting $F_\theta = a^2 e^{jka} f_\theta$

$$f_\theta = \int_0^{\pi/2} d\vartheta' \left\{ \left[H_0 \cos\theta + \sqrt{\frac{\epsilon}{\mu}} E_0 \right] \cdot \cos\Phi \cos\vartheta' \sin\vartheta' e^{-jka2\sin^2\theta/2 \cos\vartheta'} \cdot \int_0^{2\pi} e^{jka \sin\theta \sin\vartheta' \cos(\Phi - \phi')} d\phi' \right\} + \int_0^{\pi/2} d\vartheta' H_0 \sin\theta \sin^2\vartheta' e^{-jka2\sin^2\theta/2 \cos\vartheta'} \cdot \int_0^{2\pi} \cos\phi' e^{+jka \sin\theta \sin\vartheta' \cos(\Phi - \phi')} d\phi. \quad (5)$$

The integral with respect to ϕ' in the first term is

$$\int_0^{2\pi} e^{jka \sin\theta \sin\vartheta' \cos(\Phi - \phi')} d\phi' = 2\pi J_0(ka \sin\theta \sin\vartheta').$$

We now specialize to the planes $\Phi = 0$ (the x, z plane) and $\Phi = \pi/2$ (the y, z plane).

The integral with respect to ϕ' in the second term of (5) is then:

Plane $\Phi=0$

$$\int_0^{2\pi} \cos\phi' e^{jka\sin\theta\sin\vartheta'\cos\phi'} d\phi' = 2\pi j J_1(ka\sin\theta\sin\vartheta')$$

Plane $\Phi=\pi/2$

$$\int_0^{2\pi} \cos\phi' e^{jka\sin\theta\sin\vartheta'\sin\phi'} d\phi' = \int_{\phi'=0}^{\phi'=2\pi} e^{jka\sin\theta\sin\vartheta' x} dx = 0 (x=\sin\phi').$$

The latter integral shows, taken together with the fact that a factor of $\cos\Phi$ appears in the first term of (5), that E_θ vanishes everywhere in the plane $\Phi=\pi/2$, as could have been predicted beforehand.

In the plane $\Phi=0$, we have:

$$f_\theta(\theta, \Phi=0) = 2\pi \int_0^{\pi/2} d\vartheta' \left[H_0\cos\theta + \sqrt{\frac{\epsilon}{\mu}} E_0 \right] \cdot \cos\vartheta' \sin\vartheta' e^{-jka2\sin^2\theta/2\cos\vartheta'} J_0(ka\sin\theta\sin\vartheta') + \sin\theta 2\pi j \int_0^{\pi/2} d\vartheta' H_0 \cdot \sin^2\vartheta' e^{-jka2\sin^2\theta/2\cos\vartheta'} J_1(ka\sin\theta\sin\vartheta').$$

If we now assume that $\vec{E}$ and $\vec{H}$ at the aperture are related as in a plane wave, we have,

letting $G_\theta = f_\theta/2\pi$,

$$u = 2ka\sin^2\frac{\theta}{2}, \quad \text{and} \quad v = ka\sin\theta;$$

$$G_\theta(\theta, \Phi=0) = (1+\cos\theta)\int_0^{\pi/2} H_0\cos\vartheta'\sin\vartheta' e^{-iu\cos\vartheta'} J_0(v\sin\vartheta') d\vartheta' + \sin\theta j \int_0^{\pi/2} H_0 \sin^2\vartheta' e^{-iu\cos\vartheta'} J_1(v\sin\vartheta') d\vartheta'. \quad (6)$$

Considering the first integral:

$$I_1 = \int_0^{\pi/2} H_0\cos\vartheta'\sin\vartheta' e^{-iu\cos\vartheta'} J_0(v\sin\vartheta') d\vartheta' = \frac{1}{v^2}\int_0^v H_0 e^{-iu}\sqrt{1-\left(\frac{x}{v}\right)^2}\, x J_0(x) dx$$

where $x = v\sin\vartheta'$.

Integrating by parts:

$$I_1 = H_0\left(\frac{\pi}{2}\right)\frac{J_1(v)}{v} - j\frac{u}{v}\int_0^{\pi/2} \sin^2\vartheta' J_1(v\sin\vartheta') e^{-iu\cos\vartheta'} H_0 d\vartheta' - \frac{1}{v}\int_0^{\pi/2} \sin\vartheta' J_1(v\sin\vartheta') e^{-iu\cos\vartheta'} (dH_0/d\vartheta') d\vartheta'. \quad (7)$$

To obtain the first term in (6), (7) has to be multiplied by $(1+\cos\theta)=2\cos^2\theta/2$. The second term in (7) becomes

$$= -j\sin\theta\int_0^{\pi/2} \sin^2\vartheta' J_1(v\sin\vartheta') e^{-iu\cos\vartheta'} H_0 d\vartheta'$$

and this just cancels the last term in (6). Hence we have:

$$G_\theta(\theta, \Phi=0) = (1+\cos\theta)\frac{J_1(v)}{v} H_0\left(\frac{\pi}{2}\right) - \frac{\cot\frac{\theta}{2}}{ka}\int_0^{\pi/2} \sin\vartheta' J_1(v\sin\vartheta') e^{-iu\cos\vartheta'} (dH_0/d\vartheta') d\vartheta'. \quad (8)$$

We now have to introduce an assumption regarding the amplitude distribution of the aperture field. The geometrical optics field in the aperture will, in general, exhibit a singularity at the edge of the lens, since the radiation pattern of the feed must be multiplied by $\sec^{1/2}\vartheta'$ to obtain the aperture field as a function of the central angle ϑ'. Therefore, unless the feed pattern falls to zero (and more rapidly than $\cos^{1/2}\vartheta'$) at 90 degrees, the field intensity at $\vartheta'=90$ degrees in the aperture becomes infinite. In practice, of course, this is not the case since the reasoning based on geometrical optics is not valid in such regions of rapidly varying field strength, and, in addition, losses in the dielectric would tend to reduce the field strength near the edge of the lens.

A very flexible aperture distribution to choose is $1+A\cos^p\vartheta'$. This means that the illumination taper is $1+A$, and the factor $H_0(\pi/2)$ appearing in the first term of (8) is just unity. By assigning various values to p and A we have considerable control over the way in which the amplitude varies over the aperture. Thus, for example, by using negative values of A we have fields which increase towards the edge of the aperture; by using positive values of A we have fields which decrease towards the edge. Different values of p determine how peaked the distribution is.

With this aperture distribution the integral in (8) becomes (apart from a factor)

$$I_p = \int_0^{\pi/2} \sin^2\vartheta' \cos^{p-1}\vartheta' J_1(v\sin\vartheta') e^{-iu\cos\vartheta'} d\vartheta'.$$

Expanding the exponential

$$I_p = \sum_{n=0}^{\infty} \frac{(-ju)^n}{\Gamma(n+1)}\int_0^{\pi/2} \sin^2\vartheta' J_1(v\sin\vartheta')\cos^{n+p-1}\vartheta' d\vartheta'.$$

The integral can be evaluated by Sonine's formula:

$$\int_0^{\pi/2} J_\mu(v\sin\vartheta')\sin^{u+1}\vartheta'\cos^{2\alpha+1}\vartheta' d\vartheta' = \frac{2^\alpha\Gamma(\alpha+1)}{v^{\alpha+1}} J_{\mu+\alpha+1}(v)$$

with $\mu=1$ $\qquad \alpha = \dfrac{n+p}{2} - 1$

$$I_p = \sum_{n=0}^{\infty} \frac{(-ju)^n}{\Gamma(n+1)} \frac{2^{n+p/2-1}\Gamma\left(\frac{n+p}{2}\right)}{v^{n+p/2}} J_{(n+p)/2+1}(v)$$

magnetic currents $\hat{n}\times\vec{H}$ and $-\hat{n}\times\vec{E}$ flowing on the surface. Since there are no sources in the volume, this contribution must be zero. Hence the contribution from the circular base (which corresponds to the circular aperture) must equal the contribution from the spherical cap (which corresponds to the lens).

$$I_p = \frac{1}{2}\left(\frac{2}{v}\right)^{p/2} \sum_{m=0}^{\infty} (-1)^m \left(\frac{u^2}{v}\right)^m 2^m \frac{\Gamma\left(m+\frac{p}{2}\right)}{\Gamma(2m+1)} J_{m+1+(p/2)}(v)$$

$$- j\frac{u}{\sqrt{v}}\frac{1}{\sqrt{2}}\left(\frac{2}{v}\right)^{p/2} \sum_{m=0}^{\infty} (-1)^m \left(\frac{u^2}{v}\right)^m 2^m \frac{\Gamma\left(m+\frac{1}{2}+\frac{p}{2}\right)}{\Gamma(2m+2)} J_{m+(3/2)+(p/2)}(v). \quad (9)$$

Substituting this into (8), and putting in the factors which were omitted previously, the final expression for the pattern is:

$$G_\theta(\theta, \Phi=0) = (1+\cos\theta)\frac{J_1(v)}{v} + \frac{pA}{uka}\left(\frac{2}{v}\right)^{p/2-1} \sum_{m=0}^{\infty} (-1)^m \left(\frac{u^2}{v}\right)^m \frac{2^m}{(2m)!}\Gamma\left(m+\frac{p}{2}\right) J_{m+1+(p/2)}(v)$$

$$- j\frac{pA}{ka}\left(\frac{2}{v}\right)^{(p-1)/2} \sum_{m=0}^{\infty} (-1)^m \left(\frac{u^2}{v}\right)^m \frac{2^m}{(2m+1)!}\Gamma\left(m+\frac{p}{2}+\frac{1}{2}\right) J_{m+(3/2)+(p/2)}(v). \quad (10)$$

If we now return to (4), and proceed as before, using the fact that $\vec{E}$ and $\vec{H}$ in the aperture are related by the factor $\sqrt{\mu/\epsilon}$, we obtain exactly the same expression for G_Φ. Thus the patterns in the two principal planes are identical.

As $A\to 0$ (constant illumination) the pattern approaches that of a circular aperture having the same radius as the sphere. That this must be so can be seen from the following argument. We consider a plane wave traveling in the z-direction, and excise a hemispherical volume (oriented as shown in Fig. 1) from the field. The contribution from the sources inside this volume to the field may be computed from the fictitious electric and

Particular Aperture Distributions

Let us now consider the explicit expressions obtained for specific values of p. For $p=1$ we can transform (9) by means of the duplication formula for the Gamma function:

$$\frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)}{\Gamma(n+1)} = \frac{\sqrt{\pi}}{2^n\Gamma\left(\frac{n}{2}+1\right)}$$

which leads to the following expression for the pattern:

$$G_\theta = (1+\cos\theta)\frac{J_1(v)}{v} + \frac{A}{ka}\frac{v}{u}\sum_{m=0}^{\infty}(-1)^m\left(\frac{u^2}{v}\right)^m \frac{1}{2^m\Gamma(m+1)}\sqrt{\frac{\pi}{2v}}J_{m+(3/2)}(v)$$

$$- j\frac{A}{ka}\frac{\sqrt{\pi}}{2}\frac{u}{v}\sum_{m=0}^{\infty}(-1)^m\left(\frac{u^2}{v}\right)^m \frac{1}{2^m}\frac{1}{\Gamma\left(m+\frac{3}{2}\right)}J_{m+2}(v)$$

or:

$$G_\theta = (1+\cos\theta)\frac{J_1(v)}{v} + \frac{A}{ka}\frac{v}{u}\left[\sqrt{\frac{\pi}{2v}}J_{3/2}(v) - \frac{1}{2}\left(\frac{u^2}{v}\right)\sqrt{\frac{\pi}{2v}}J_{5/2}(v)\right.$$

$$\left. + \frac{1}{2\cdot 4}\left(\frac{u^2}{v}\right)^2\sqrt{\frac{\pi}{2v}}J_{7/2}(v) - \frac{1}{2\cdot 4\cdot 6}\left(\frac{u^2}{v}\right)^3\sqrt{\frac{\pi}{2v}}J_{9/2}(v) + \frac{1}{2\cdot 4\cdot 6\cdot 8}\left(\frac{u^2}{v}\right)^4\sqrt{\frac{\pi}{2v}}J_{11/2}(v) - \cdots\right]$$

$$- j\frac{A}{ka}\left[J_2(v) - \frac{1}{3}\left(\frac{u^2}{v}\right)J_3(v) + \frac{1}{3\cdot 5}\left(\frac{u^2}{v}\right)^2 J_4(v) - \frac{1}{3\cdot 5\cdot 7}\left(\frac{u^2}{v}\right)^3 J_5(v) + \frac{1}{3\cdot 5\cdot 7\cdot 9}\left(\frac{u^2}{v}\right)^4 J_6(v) - \cdots\right].$$

For the range of θ of interest in practice, the series converges extremely rapidly [2]. For the main beam and first few side lobes it is only necessary to take the first two real terms, and none of the imaginary terms, except right in the vicinity of the nulls where one or two additional terms in each series may be required for very high accuracy.

Although the radius of the lens appears explicitly in the formulas, it can be eliminated for any lens likely to be of interest in practice. This makes it possible to write everything in terms of the single variable $v=ka \sin\theta$, which in turn makes it possible to derive general formulas for quantities which characterize the performance of the lens, independent of the lens size.

To do this, we note that $u=ka-\sqrt{(ka)^2-v^2}$ may be expanded by the binomial theorem to read $u=v^2/2ka+v^4/\delta(ka)^3+\cdots$. The second term is less than about 1 per cent of the first for values of ka and v which are of interest in practice ($ka>\sim 40$, $v<\sim 10$ for main beam and first side lobe.) Hence, to derive general formulas for beam width, first side lobe level, and aperture efficiency, we can write $u=v^2/2ka$.

For $p=1$, we have

$$G_\theta \cong \frac{2J_1(v)}{v}+\frac{2A}{v}\left[\sqrt{\frac{\pi}{2v}}\,J_{3/2}(v)\right].$$

Beamwidth Formulas

To find formulas for the beamwidth, we have to solve the equation

$$G_\theta(v) = 0.7071 G_\theta(0).$$

At $v=0$ the value of G_θ may be obtained by noting that $J_{3/2}$ may be written in closed form as

$$J_{3/2}(v) = \sqrt{\frac{2}{\pi v}}\left(\frac{\sin v}{v}-\cos v\right),$$

hence

$$G_\theta(v) \cong \frac{2J_1(v)}{v}+\frac{2A}{v^2}\left(\frac{\sin v}{v}-\cos v\right). \qquad (11)$$

Expanding $\sin v$ and $\cos v$, the expression for $G_\theta(v)$ in the neighborhood of $v=0$ becomes:

$$G_\theta(v) \cong 1+2A\frac{1}{v^2}\left\{1-\frac{v^2}{6}+\cdots-1+\frac{v^2}{2}-\cdots\right\}$$

$$G_\theta(0) = 1+\frac{2A}{3}. \qquad (11a)$$

Therefore the transcendental equation to be solved for the beamwidths is:

$$J_1(v)+A\sqrt{\frac{\pi}{2v}}\,J_{3/2}(v) = 0.7071\left(0.5+\frac{A}{3}\right)v. \qquad (12)$$

(The quantity $\sqrt{(\pi/2v)}J_{n+1/2}(v)$ has been retained throughout since it is tabulated as $Q_n(v)$ in the WPA table of "Spherical Bessel Functions.")

Solution of the transcendental equation for various values of A results in the following beamwidth formulas:

Taper	*Half-Power Beamwidth*
-20 *db*	$\theta_{1/2} = 25.7\dfrac{\lambda^\circ}{a}$
-10 *db*	$27.2\dfrac{\lambda^\circ}{a}$
0 *db*	$29.6\dfrac{\lambda^\circ}{a}$
10 *db*	$31.6\dfrac{\lambda^\circ}{a}$
20 *db*	$32.5\dfrac{\lambda^\circ}{a}$

(Positive values of taper refer to a *decrease* in field intensity at the edge of the aperture.)

Side Lobe Level

To compute the side lobe level as a function of illumination taper, we first differentiate (11) with respect to v, and set the result equal to zero:

$$\frac{d}{dv}\left\{2\frac{J_1(v)}{v}+\frac{2A}{v^2}\left(\frac{\sin v}{v}-\cos v\right)\right\} = 0.$$

This results in the expression:

$$A = \frac{\frac{1}{3}v^2J_2(v)}{\cos v-\dfrac{\sin v}{v}+\dfrac{v\sin v}{3}}.$$

We would now like to solve this equation for v as a function of A, and calculate the side lobe position for various values of A. Then we could substitute the side lobe position in (11) to obtain the side lobe level. However, since we cannot solve for v as a function of A, we consider the two equations

$$A = \frac{\frac{1}{3}v^2J_2(v)}{\cos v-\dfrac{\sin v}{v}+\dfrac{v\sin v}{3}} \qquad (13)$$

and

$$L = -20\log\left|\frac{1+\dfrac{2A}{3}}{\dfrac{2J_1(v)}{v}+\dfrac{2A}{v}\left[\sqrt{\dfrac{\pi}{2v}}\,J_{3/2}(v)\right]}\right|$$

as a pair of parametric equations. By substituting different values of v into these equations, we obtain corresponding values of taper $(1+A)$ and side lobe level (L). (The equation for L has been normalized to zero at the peak of the main beam.)

Values of v from about 4.6 up to about 6 will give all tapers of practical interest. This information on the ap-

proximate range of v may be determined from a plot of the field (Fig. 2) for no taper, which shows that v must be of the order of 5.2; v is then increased from this point until the denominator in the expression for A changes sign, since the point where it passes through zero corresponds to infinite taper, *i.e.*, zero field at the edge of the lens. It is then also decreased until $A=-1$, since this includes all cases where the field increases in going from the center of the aperture to the edge.

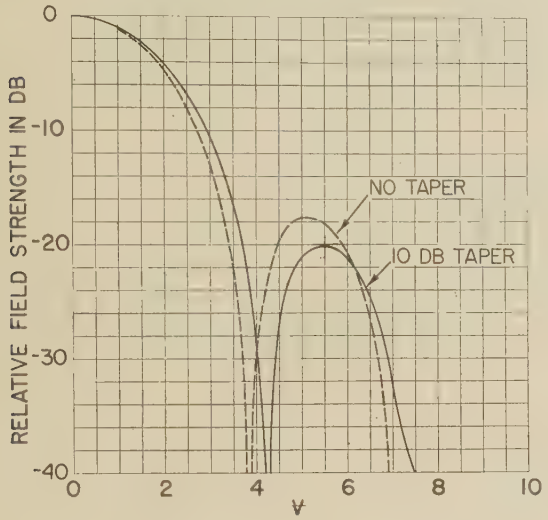


Fig. 2

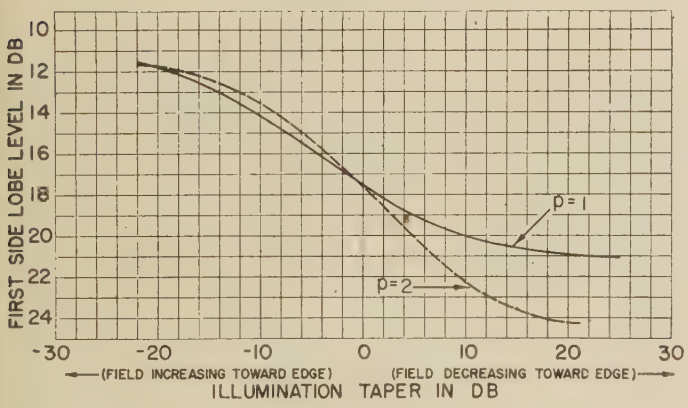


Fig. 3

The curve of side lobe level versus illumination taper obtained from these parametric equations is given in Fig. 3.

It is also of interest to examine the distribution $\cos\vartheta'$, where the field drops to zero at the edge of the aperture. The position of the first side lobe is obtained from the condition that the denominator of (13) vanish. This leads to the transcendental equation

$$\tan v = \frac{-3v}{v^2-3}$$

which has the solution $v=5.764$. The side lobe level is then obtained from the second of (13) by substituting $A=\infty$, and $v=5.764$

$$L = -20\log\left|\frac{\frac{v}{3}}{\sqrt{\frac{\pi}{2v}}J_{3/2}(v)}\right| = -21.3 \text{ db.}$$

If we now consider the field produced by the distribution $1+A\cos^2\vartheta'$, which is a more peaked distribution, we have to substitute $p=2$ in (10). Again we find it necessary to retain only the first two terms in the real part of the series, and we can neglect the imaginary part of the series. The expression for the field in this case is

$$G_\theta \cong \frac{2J_1(v)}{v} + \frac{4A}{v^2}J_2(v).$$

To evaluate this at $v=0$, one can apply l'Hospital's rule twice, or else use the series expansion for $J_n(v)$ resulting. in

$$G_\theta(0) = 1 + \frac{A}{2}. \qquad (13a)$$

To determine the beamwidth, we have to solve the equation

$$\frac{2J_1(v)}{v} + \frac{4A}{v^2}J_2(v) = 0.7071\left(1+\frac{A}{2}\right)$$

which can be written in terms of the Λ functions tabulated in Jahnke and Emde:

$$\Lambda_1 + \frac{A}{2}\Lambda_2 = 0.7071\left(1+\frac{A}{2}\right).$$

For a 10 db taper ($A=2.17$) the root of this equation is approximately 1.8. This leads to the formula for the half-power beamwidth

$$\theta_{1/2} = 32.8\frac{\lambda^\circ}{a} \quad (10\,\text{db taper}, 1+A\cos^2\vartheta' \text{ dist.}).$$

The side lobe level vs illumination taper is again determined in the same way by a pair of parametric equations:

$$A = \frac{2J_1(v) - vJ_0(v)}{J_1(v) + \frac{4J_2(v)}{v} - J_3(v)}$$

$$L = -20\log\left|\frac{v\left(1+\frac{A}{2}\right)}{2J_1(v) + 4A\frac{J_2(v)}{v}}\right|.$$

A plot of these equations is given in Fig. 3. More peaked distributions may be discussed by substituting higher values of p in (10).

Gain

The gain of the lens may be computed from the expression

$$\text{Gain} = \frac{4\pi \frac{1}{2} R^2 \sqrt{\frac{\epsilon}{\mu}} \mid E(0) \mid^2}{P_T} \tag{14}$$

where $|E|$ in the numerator is evaluated at $\theta = 0$, and P_T is the total power radiated. P_T may be computed by integrating the Poynting vector over the aperture of the lens; the result is

$$P_T = \frac{1}{2} a^2 2\pi \sqrt{\frac{\mu}{\epsilon}} \left[\frac{1}{2} + \frac{2A}{p+2} + \frac{A^2}{2p+2} \right].$$

Using the expression for $E_\theta(0)$:

$$E_\theta(0) = \frac{-j\omega\mu}{4\pi} \frac{e^{-jkR}}{R} a^2 e^{jka} 2\pi G_\theta(0),$$

the numerator of (14) becomes

$$4\pi \frac{1}{2} \sqrt{\frac{\epsilon}{\mu}} \left[\frac{\omega^2\mu^2}{4} a^4 \mid G_\theta(a) \mid^2 \right].$$

Combining this with the expression for P_T, and rearranging:

$$\text{Gain} = \frac{4\pi}{\lambda^2} (\pi a^2) \frac{\mid G_\theta(0) \mid^2}{\left[1 + \frac{4A}{p+2} + \frac{A^2}{p+1} \right]}.$$

Thus the efficiency of the lens, η, compared to the equivalent uniformly illuminated circular aperture, is

$$\eta_{pA} = \frac{\mid G_\theta(0) \mid^2}{1 + \frac{4A}{p+2} + \frac{A^2}{p+1}}.$$

For $p=1$, *i.e.*, for the aperture distribution $1 + A \cos \vartheta'$, we have from (11a) that $|G_\theta(0)| = 1 + \frac{2}{3}A$. Hence

$$\eta_{1A} = \frac{(1 + \frac{2}{3}A)^2}{1 + \frac{4A}{3} + \frac{A^2}{2}}.$$

This is plotted in Fig. 4. For $p=2$, $|G_\theta(0)| = 1 + (A/2)$ from (13a). Hence

$$\eta_{2A} = \frac{\left(1 + \frac{A}{2}\right)^2}{1 + A + \frac{A^2}{3}}.$$

This is also plotted in Fig. 4.

For the distribution $\cos \vartheta'$, we can obtain the efficiency from η_{1A} by allowing A to become infinite:

$$\eta_{1\infty} = \frac{(\frac{2}{3})^2}{\frac{1}{2}} = \frac{8}{9} = 0.889.$$

For the distribution $\cos^2 \vartheta'$, we allow A to become infinite in η_{2A}:

$$\eta_{2\infty} = \frac{(\frac{1}{2})^2}{\frac{1}{3}} = \frac{3}{4} = 0.75.$$

Again, efficiency with more peaked distributions may be calculated by proceeding in the same way from (10).

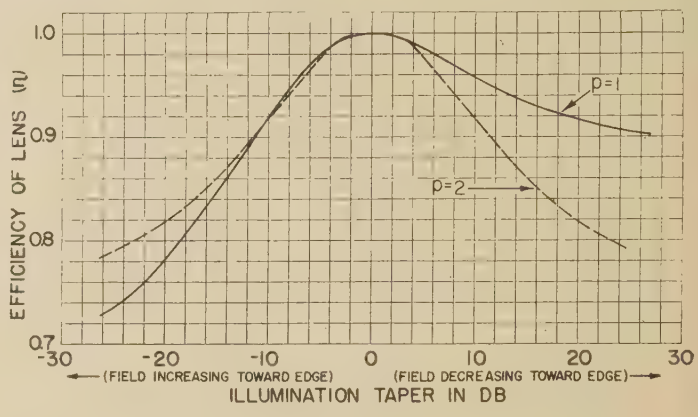


Fig. 4

The spherical Luneberg lens has a lower sidelobe level and a somewhat wider beamwidth than the corresponding two dimensional lens having the same illumination taper. This is due principally to the effective illumination taper produced by the circular aperture shape.

Conclusion

Although the precise aperture distribution which will be obtained with a given primary feed pattern cannot be predicted exactly on the basis of geometrical optics, these results indicate that satisfactory far field patterns are obtainable with almost any reasonable amplitude taper. In addition they provide information as to the magnitude of the changes required in the amplitude distribution to produce significant improvements in side lobe level, beam width, and gain, and make it possible to arrive at a compromise distribution which will result in satisfactory values for each of these three parameters.

Bibliography

[1] Luneberg, R. K., "Mathematical Theory of Optics." Brown University.
[2] An expression giving an upper limit to the error incurred by omitting all but the first N terms has been derived; it depends essentially on comparing the series in question with an exponential series.
[3] Rinehart, R. F., "A Solution of the Problem of Rapid Scanning for Radar Antennas." *Journal of Applied Physics* (September, 1948).
[4] Rinehart, R. F., "A Family of Designs for Rapid Scanning Radar Antennas." PROCEEDINGS OF THE IRE (June, 1952).
[5] Peeler, G. D. M., Archer, D. H., and Kelleher, K. S., "A Microwave Luneberg Lens." IRE Convention (March, 1952).
[6] Jasik, H., "The Electromagnetic Theory of the Luneberg Lens." Air Force Cambridge Research Center Report (November, 1954).

Antenna Pattern Distortion by Dielectric Sheets*

J. H. RICHMOND†

Summary—**The radiation pattern of an antenna covered with a radome can be determined from the following data: The antenna Fresnel-Zone fields in the absence of the radome, and the fields inside the radome produced by a plane-wave incident from an external source.**

In certain cases the pattern calculations may be simplified without excessive loss of accuracy by using optical theory to describe approximately the plane-wave transmission through the radome, and neglecting scattering by the antenna in determining the fields inside the radome from the external plane-wave source.

Introduction

IN DESIGNING a radome it is often necessary to calculate the far-field pattern of an antenna covered with the radome. Such calculations are generally based on given data on the Fresnel fields of the antenna transmitting without a radome. Ray-tracing methods are used to describe approximately the transmission of these fields through the dielectric radome. It will be shown that the problem can be reformulated in a manner which requires a knowledge of transmission through the radome only for a simple plane wave, rather than for the complicated Fresnel-Zone waves of the antenna.

Theory

Consider the problem of determining the far-field pattern of antenna 1 shown in Fig. 1. Let E_1 and H_1 denote the fields of antenna 1 when transmitting. Choose a reference surface S which must be either an infinite plane or any surface which completely encloses antenna 1. In computing the fields of antenna 1 on the source-free side of S, antenna 1 can be replaced by its equivalent electric and magnetic current distributions J_1 and K_1 on S, where

$$J_1 = n_1 \times H_1, \tag{1a}$$

and

$$K_1 = E_1 \times n_1, \tag{1b}$$

where n_1 is a unit vector in the direction of the outward-drawn normal on S. Two vector potentials A and F can now be defined in terms of J_1 and K_1 in the usual manner,[1] and the field intensity can be expressed in terms of A and F. If the entire region on the source-free side of S is occupied by a homogeneous medium, the potentials and the electric field intensity in this region are given by:[1]

$$A = \frac{\mu}{4\pi}\int\int_S J \frac{e^{-jkr}}{r}\, ds, \tag{2a}$$

$$F = \frac{\epsilon}{4\pi}\int\int_S K \frac{e^{-jkr}}{r}\, ds, \tag{2b}$$

and

$$E = -\frac{1}{\epsilon}\nabla \times F - \frac{j}{\mu\omega\epsilon}\nabla \times \nabla \times A. \tag{2c}$$

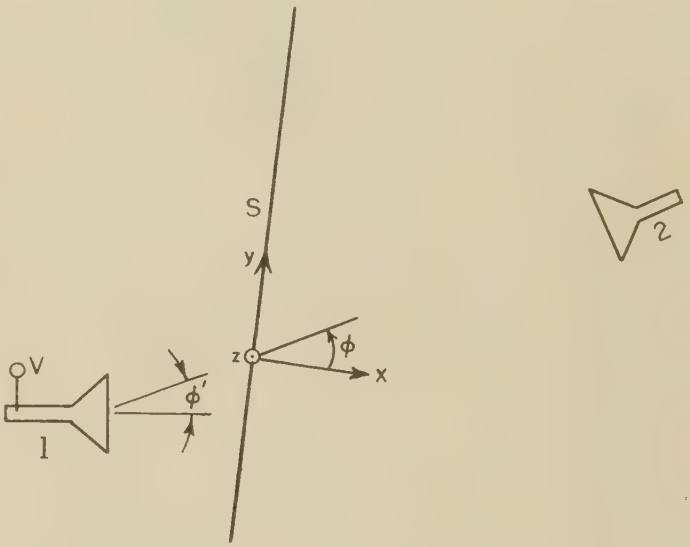


Fig. 1—Coordinates used in radiation pattern analysis.

The expressions for A and F become much more complicated for the case of an inhomogeneous region. Therefore, radome calculations are usually divided into two steps. First, an auxiliary reference surface S'' is chosen between antenna 1 and the radome. The unperturbed fields of antenna 1 on S'' are used to obtain an approximation for the fields on surface S which is outside the radome. Eqs. 2 are then used to compute the far-field pattern. This type of analysis, which may be called the conventional approach to the radome problem at present, requires a mathematical description of the transmission of energy from antenna 1 through the radome to the external reference surface.

The problem may be reformulated by introducing a distant transmitting antenna, antenna 2, as in Fig. 1. The far-field pattern of antenna 1 is given by the voltage V induced at its terminals as antenna 2 is moved in a circular path centered at antenna 1. Antenna 2 is always aimed at antenna 1 during this motion. Let E_2 and H_2 denote the electric and magnetic field intensities on S due to antenna 2 transmitting *in the presence of antenna 1*. Antenna 2 can be replaced by its equivalent electric

* Manuscript received by the PGAP, January 2, 1956. The research reported in this paper was sponsored by the Air Res. and Dev. Com., Wright Air Dev. Center, Wright-Patterson AFB, Ohio.

† Antenna Lab., Dept. Elec. Engrng., Ohio State Univ., Columbus, Ohio.

[1] A. B. Bronwell and R. E. Beam, "Theory and Application of Microwaves," 1st ed., McGraw-Hill Book Co., New York, N. Y., pp. 451–453; 1947.

and magnetic current distributions J_2 and K_2 on S, where

$$J_2 = n_2 \times H_2, \tag{3a}$$

and

$$K_2 = E_2 \times n_2, \tag{3b}$$

where n_2 is the unit vector normal to S in the direction of the region occupied by antenna 1. By an application of the Lorentz reciprocal theorem, Kouyoumjian[2] has shown that

$$V \propto \iint_S (J_2 \cdot E_1 - K_2 \cdot H_1) ds. \tag{4}$$

As will be shown, the additional information introduced by the application of the reciprocity theorem makes it possible to simplify the radome analysis by strategic use of (4) in place of the conventional approach.

If S is chosen to be a plane, (4) simplifies to

$$V \propto \iint_S J_2 \cdot E_1 ds. \tag{5}$$

For far-field calculations, antenna 2 is at a great distance from antenna 1, the field on S due to antenna 2 is a plane wave (if scattering from antenna 1 may be neglected), and (5) reduces to the usual expression for the far-field pattern of antenna 1 in terms of its aperture fields.

To illustrate the application of (4) to the radome problem, consider the far-field pattern of antenna 1 in the presence of an infinite plane dielectric sheet as in Fig. 2. Choose as a reference the surface of the dielectric sheet facing antenna 1. Now to use (4) or (5) it is necessary to know: the fields on S of antenna 2 transmitting through the dielectric sheet in the presence of antenna 1, and the unperturbed fields on S of antenna 1 when it is transmitting. Neglecting the scattered fields of antenna 1, the fields of antenna 2 on S are found by employing the plane-wave plane-sheet transmission coefficient T. Using primes to denote quantities in the presence of the dielectric and the absence of primes to denote the corresponding quantities without the dielectric,

$$H_2' \approx TH_2, \tag{6a}$$

and

$$J_2' \approx TJ_2. \tag{6b}$$

From (5),

$$V' \propto \iint_S J_2' \cdot E_1 ds. \tag{7}$$

[2] R. G. Kouyoumjian, "The Calculation of the Echo Areas of Perfectly Conducting Objects by the Variational Method," Report 444-13, Contract DA 36-039 sc 5506, Ohio State University Res. Found., pp. 29 and 80–84; Nov. 15, 1953.

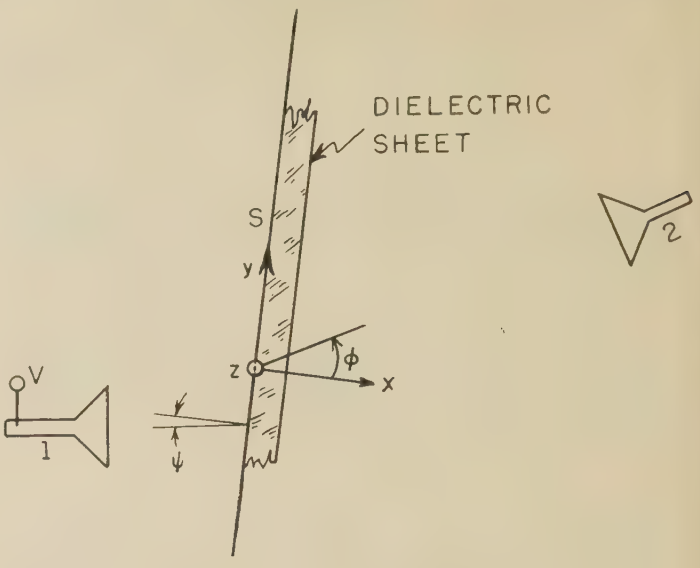


Fig. 2—Transmission through an infinite plane dielectric sheet.

From (6) and (7), since T is independent of position on S,

$$V' \approx T \iint_S J_2 \cdot E_1 ds = TV, \tag{8a}$$

or

$$V'(\phi) \approx T(\phi) V(\phi). \tag{8b}$$

That is, knowing the unperturbed far-field pattern V of antenna 1, its pattern V' in the presence of an infinite plane dielectric sheet is found merely by multiplication by the plane-wave plane-sheet transmission coefficient for the polarization and angle of arrival of the plane wave coming from antenna 2. This contrasts with the usual optical approximations in which the transmission coefficient for the angle of incidence ψ (Fig. 2) of rays from antenna 1 is employed.

Suppose the dielectric surface facing antenna 2 had been chosen as the reference surface. The dielectric sheet would then be considered a part of antenna 1 and it would be necessary to know the fields of antenna 1 on this surface in the presence of the dielectric sheet. Thus, a surface passing between the radome and antenna 1, or perhaps coinciding with the dielectric surface facing antenna 1, is preferred since transmission through a dielectric body is most easily calculated for plane waves.

Eqs. (8) are approximate in that they neglect the waves scattered by antenna 1. To determine the error involved in this approximation, several examples of transmission through infinite plane sheets and finite plane sheets will now be considered.

Examples

Fig. 3 shows the measured far-field H plane patterns of a pyramidal horn at 9,400 mc with a polystyrene ($\epsilon = 2.53$) plane sheet held at an angle of 40° across the

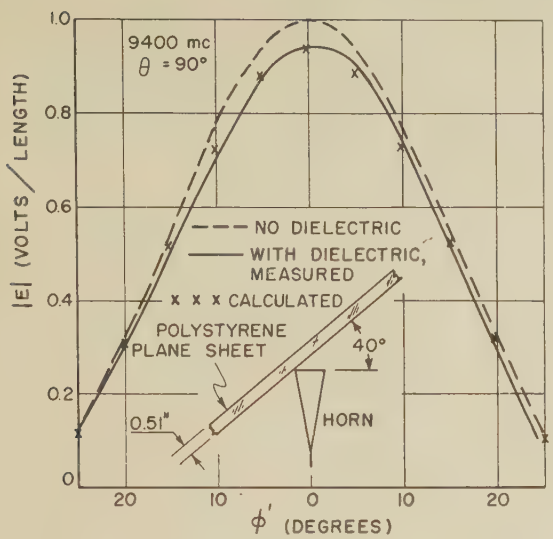


Fig. 3—Far-field pattern of horn trahsmitting thrausmitting a plane polystyrene sheet.

horn aperture. The unperturbed pattern is also shown for reference. The calculated points shown were obtained using (8b). The dielectric sheet was 0.51 inch thick and was large enough so that its edges were well outside the region illuminated by the horn. The horn had a length of 7.8 inches, flare angles of 20°, and a beam width of 23° to the half-power points. The waves arriving from antenna 2 were polarized perpendicular to the plane of incidence on the dielectric sheet. Hence, the transmission coefficient for perpendicular polarization was employed in the calculations.

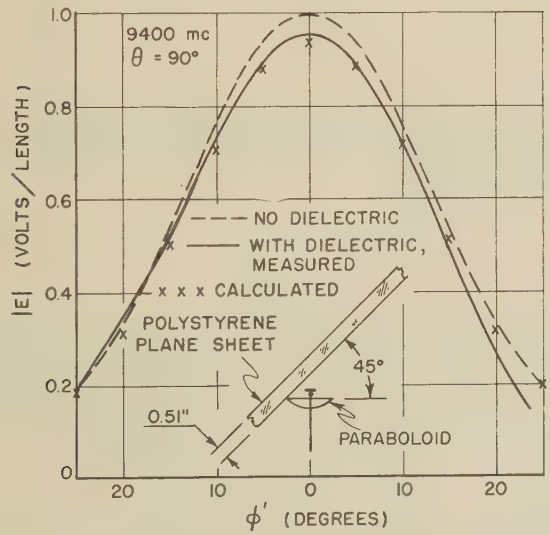


Fig. 4—Far-field pattern of paraboloid transmitting through a plane polystyrene sheet.

Fig. 4 shows similar results for a paraboloidal antenna with the same polystyrene sheet at 45°. The paraboloid had a diameter of 4 inches and a beam width of 23°.

One more example using the horn is shown in Fig. 5 for a fiberglas sheet having a thickness of 0.382 inch and a dielectric constant of 4. This sheet had a lower transmission coefficient than the polystyrene sheet and, therefore, produced a greater reduction in the far-field amplitude. The fiberglas and polystyrene sheets were each a little over one-half wavelength thick in terms of the wavelength in the dielectric.

To get closer to the radome problem, calculations were also carried out for plane dielectric sheets covering only a part of the antenna aperture. A reference surface S was chosen on an infinite plane coinciding with the dielectric surface facing antenna 1. The fields of antenna 1 were measured on S in the absence of the dielectric sheet. Eq. (4) or (5) requires a knowledge of the

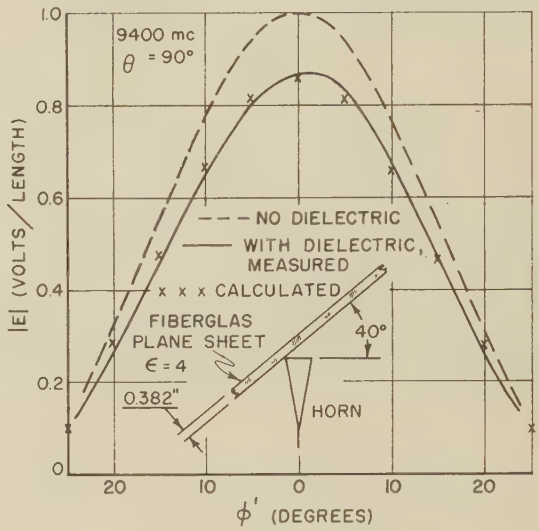


Fig. 5—Far-field pattern of horn transmitting through a plane fiberglas sheet.

fields on S due to antenna 2 radiating in the presence of the dielectric. Although at points remote from the edge of the sheet the fields are given by plane-wave plane-sheet theory, no solution appears to be available for the fields near the edge. As a first approximation, the fields were assumed to be the incident plane wave up to the edge of the dielectric sheet, and were assumed to be given by plane-wave plane-sheet theory on the portion of S covered with dielectric. The results for this approximate calculation, using (5), are shown in Figs. 6 and 7 for a polystyrene sheet 0.51 inch thick covering one-half of the horn aperture at angles of 40° and 0°, respectively.

Conclusion

Knowing the unperturbed fields of an antenna, it is often desired to calculate the radiation pattern of the antenna covered with a radome. Generally, optical ray-tracing methods are employed to describe approximately the transmission of the antenna Fresnel-Zone fields outward through the radome. By applying the reciprocity theorem, however, this problem can be

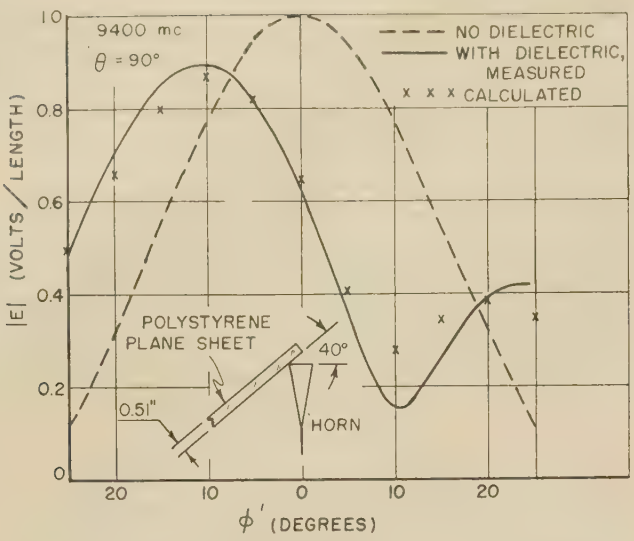


Fig. 6—Far-field pattern of horn with aperture partially covered with a plane polystyrene sheet.

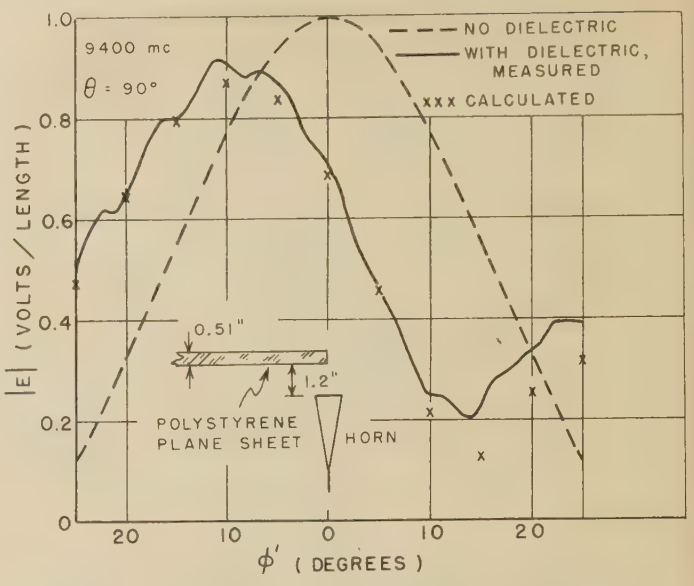


Fig. 7—Far-field pattern of horn with aperture partially covered with a plane polystyrene sheet.

traded for that of computing the fields inside the radome for a plane wave arriving from an external source. This problem is simplified if scattering by the antenna can be neglected. To determine the error involved in neglecting this scattering, measurements and calculations were made of the patterns of small horn and paraboloidal antennas covered with various infinite and finite plane dielectric sheets tilted at various angles from the antenna aperture. For the examples which were considered, the scattering error was quite negligible and satisfactory results were obtained by using an optical approximation to describe plane-wave transmission through the finite plane sheets.

Multiple Scattering by Randomly Distributed Obstacles —Methods of Solution*

C. M. CHU† AND S. W. CHURCHILL‡

Summary—**Methods of solution of problems of multiple scattering of electromagnetic radiation by randomly distributed obstacles are evaluated with respect to generality, accuracy, and suitability for numerical calculations. Exact solutions are found to be limited to infinite regions and are difficult to apply for anisotropic scattering. Approximate solutions of the transport equation are limited to simple regions and are very complicated for anisotropic scattering. Approximate representation of the scattering process with the classical diffusion equation permits solution for almost all conditions, but the solution is inaccurate near the boundaries of the region of scattering. Approximate representation of the intensity by a number of discrete components permits solution for many geometries. Several two-component representations have been developed which yield more accurate results than diffusion theory for anisotropic scattering. A six-component model appears to be considerably more accurate for most geometries but requires machine computations.**

* Manuscript received by the PGAP August 13, 1955; revised manuscript received January 16, 1956. Presented by J. F. Shea before the Symposium on Electromagnetic Wave Theory, sponsored by Commission VI of URSI and the University of Mich., Ann Arbor, Mich., June 20–25, 1955. This paper reports work done through the Engineering Research Institute of the University of Mich. under contract with the Chemical Corps, U. S. Army, Washington, D. C.

† Engineering Research Institute, University of Mich., Ann Arbor, Mich.

‡ Dept. of Chemical and Metallurgical Engrg., University of Mich., Ann Arbor, Mich.

Introduction

THIS PAPER presents the results of an investigation of methods of solution of problems of multiple scattering of electromagnetic radiation by randomly distributed obstacles. Proposed methods of solution are evaluated with respect to generality and to practicality for numerical calculations. Exact and approximate methods are both considered. Attention has been given to obstacles of uniform size, but results can be extended to dispersions of known size distribution.

Exact solutions have been derived for the scattering of radiation by a single obstacle of simple geometry such

as a sphere, ellipsoid, or cylindrical rod. For more complicated shapes, the amount and angular distribution of scattered power must be found by approximate methods or experimental means. For limited paths of transmission such that rescattering is negligible, the transmission and angular distribution of scattered power can be described by an exponential decay law. For sufficient lengths of transmission, secondary and higher orders of scattering become appreciable and description of the process is much more difficult.

The importance of multiple scattering has been recognized in the transmission of neutrons through dense material, of stellar radiation through dispersed matter, of light through fogs, of heat through fibrous insulation and microwaves through the atmosphere. Previous work has consequently appeared over a wide range of the scientific literature and has had diverse objectives. Most attention in the literature has been given to isotropic scattering within infinite or semi-infinite regions because such conditions are more receptive to mathematical analysis. In actual applications, the angular distribution of scattered power is generally quite anisotropic and in many practical problems the assumption of an infinite medium is not justifiable. Principal attention in this study is accordingly given to methods suitable for anisotropic scattering and finite regions.

Mathematical Formulation

Single scattering can be characterized by the following parameters:

1) σ_s, the effective cross section of an obstacle for scattering;
2) σ_t, the effective cross section of an obstacle for interaction; and
3) $f_1(\vec{\Omega}, \vec{\Omega}')$, the angular distribution function for single scattering, which gives the fraction of energy scattered from a beam with the direction, $\vec{\Omega}$, into a unit steradian in the direction, $\vec{\Omega}'$.

The cross sections and angular distribution function are functions of the index of refraction, the dimensions of the obstacle and the wavelength. For spherical obstacles the angular distribution is symmetrical about the direction of the incident radiation and may be expressed in terms of a single angle, θ, measured from the incident direction; *i.e.*, $f_1(\vec{\Omega}, \vec{\Omega}')$ reduces to $f_1(\theta)$. The combinations, $\sigma_a = \sigma_t - \sigma_s$, called the cross section for absorption, and $\omega = \sigma_s/\sigma_t$, called the albedo for single scattering, are often used for convenience.

Multiple scattering can be characterized by the above parameters for single scattering together with N, the number of obstacles per unit volume. The combinations, $1/N\sigma_s$, $1/N\sigma_a$, and $1/N\sigma_t$ called the mean free paths for scattering, absorption, and interception, respectively are also commonly employed.

The dependent variable in multiple scattering is the specific intensity, $i(\vec{R}, \vec{\Omega})$. The specific intensity is defined as the energy intercepted per unit time by a unit area at $\vec{R}$, per unit steradian from the direction $\vec{\Omega}$ normal to the area, where $\vec{R}$ is the radial vector from the origin of the coordinate system and $\vec{\Omega}$ is a unit vector representing direction. In general, $i(\vec{R}, \vec{\Omega})$ is a function of five variables, three for position and two for direction.

Multiple scattering can be formulated from either of two points of view:

1) In the random-walk or Lagrangian point of view, the energy in a wave is considered to be in the form of discrete packets (quanta). The specific intensity, $i(\vec{R}, \vec{\Omega})$ is then interpreted as the density of these packets at $\vec{R}$ traveling in the direction $\vec{\Omega}$. The mechanism of single scattering can be considered as the collision of an energy packet with an obstacle. The probability of a packet traveling in the direction, $\vec{\Omega}$, being scattered into $d\Omega'$ by a collision is $f_1(\vec{\Omega}, \vec{\Omega}')d\Omega'$, while the probability of a packet being scattered between x and $x+dx$ is $N\sigma_s e^{-N\sigma_t x}dx$, where x is the distance traveled into the dispersion or after a collision. From these two elementary probabilities, it is possible to derive an expression for the probability of finding an energy packet in a unit volume at $\vec{R}$ traveling in a unit steradian in the direction, $\vec{\Omega}$, after one, two . . . , or any number of collisions, and hence an expression for the specific intensity $i(\vec{R}, \vec{\Omega})$. Exact solutions have been developed only for an infinite region of scattering, but approximate solutions are possible for any conditions by the Monte Carlo technique.
2) In the continuous or Eulerian point of view, the scattering region is considered to be a macroscopically homogeneous media. An energy balance over a differential volume of unit cross section and length $dS_{\vec{\Omega}}$ along $\vec{\Omega}$ yields the following integro-differential equation, generally known as the transport equation:

$$\frac{di(\vec{R}-\vec{\Omega})}{dS_{\vec{\Omega}}} = +N\sigma_s \int_{\Omega'} i(\vec{R}, \vec{\Omega}')f_1(\vec{\Omega}, \vec{\Omega}')d - N\sigma_t i(\vec{R}, \vec{\Omega}). \quad (1)$$

net rate of increase in intensity in the direction $\vec{\Omega}$	rate of increase of intensity due to scattering into the direction $\vec{\Omega}$	rate of decrease of intensity in the direction $\vec{\Omega}$ due to interception

The integral term in 1 thus introduces the contribution of multiple scattering. Exact solutions of the transport equation by the Laplace transform and Fourier transforms have been postulated. Approximate solutions have been proposed by quadrature, spherical harmonics, and reiteration.

Numerous approximate representations have also been proposed for the multiple-scattering problem. The transport equation can be approximated by the classical equation of diffusion. Representation of the specific intensity by a number of discrete components also leads to approximate equations which can be solved more readily than the transport equation.

Solution by Random Walk

Exact Solutions

Exact solutions for an infinite region of scattering have been developed by Grosjean [1, 2] based on the Lagrangian point of view. The results are expressed in integral form. As an example, in the simple case of a point source in an infinite, isotropically scattering region the probability of finding an energy packet between r and $r+dr$ from the source is $\phi(r)dr$ with

$$\phi(r) = p(r) + \frac{2r}{\pi}\int_0^\infty \frac{F^2(v)}{v - F(v)} \sin rv\, dv \tag{2}$$

where

$$p(x)dx = N\sigma e^{-N\sigma_t x}\, dx \tag{3}$$

is the probability of an energy packet being scattered at a distance between x and $x+dx$ from a previous scattering, and

$$F(v) = \int_0^\infty \frac{p(x)\sin(xv)dx}{x}. \tag{4}$$

Monte Carlo Method

The increasing availability of high-speed computing machinery and current interest in neutron scattering has stimulated interest in the so-called Monte Carlo technique of numerical analysis [3]. This technique might be applied to the multiple-scattering problem as follows. The region can be divided into a number of contiguous cells of any number of shapes. Based upon the characteristic properties of the scattering region, $1/N\sigma_s$, $1/N\sigma_t$, and $f_1(\vec{\Omega}, \vec{\Omega}')$, probabilities are assigned to each cell for interception, absorption, and rescattering out of each particular face of the cell for the power entering each face. The random walk process is then carried out for a very large number of packets of energy entering the dispersion, until a statistical pattern is apparent.

The accuracy of the procedure depends on the number of cells and the number of quanta traced. In practice, the number of cells and the number of discrete directions in which scattering is allowed are limited by the storage capacity of a computer, and economically, by the computing time required to obtain a statistically representative result. Since the solution is numerical instead of analytical, the calculations must be repeated for every set of conditions.

Exact Solutions of the Transport Equation

Exact solutions of the transport equation have been worked out for only a few special cases, in all cases for infinite or semi-infinite regions and in most cases for isotropic scattering. The solutions are usually expressed in terms of definite integrals from which numerical results may be obtained by standard techniques, such as residue theory, series expansion, or numerical reiteration. In principle, these solutions can be utilized to construct solutions for finite dispersions and anisotropic scattering but the numerical procedures become too complicated for practical usage. The techniques used in two exact methods of solution are outlined below. The details can be found in the original references.

Laplace Transform Method

The Milne problem in astrophysics; *i.e.*, the determination of the radiation escaping from a star surrounded by an infinitely thick, isotropically scattering media can be solved by application of the Laplace Transform. The solution is given in terms of an "H" function which is proportional to the intensity in any direction and is defined by the equation

$$H(\mu) = 1 + \frac{1}{2}\frac{\sigma_s}{\sigma t}\mu H(\mu)\int_0^1 \frac{H(v)}{\mu + v}\, dv, \tag{5}$$

where $\mu = \cos\theta$. The method is discussed by Chandrasekhar [4] who calculated and tabulated values of the "H" function.

Fourier Transform Method

For isotropic scattering in an infinite medium involving two degrees of freedom, one for the space coordinate and one for direction, the Fourier transform of the transport equation yields a solution for the specific intensity in the form of infinite integrals. For example, for an isotropically scattering, infinite dispersion with a unit plane source at $Z=0$, the integral solution is

$$i(Z, \mu) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{A(v)e^{ivZ}}{N\sigma_t + jv\mu}, \tag{6}$$

where

$$A(v) = \frac{1}{2\left[1 - \dfrac{N\sigma_s}{2jv}\ln\dfrac{N\sigma_t + jv}{N\sigma_t - jv}\right]}. \tag{7}$$

Details are given by Glasstone and Edlund [5] and additional examples by Case, *et al.* [6].

Approximate Solutions of the Transport Equation

Quadrature Method

Quadrature has been used extensively to obtain approximate solutions of the transport equation in problems of astrophysics and neutron scattering [7]. The applications have been limited to simple geometries and isotropic scattering. The principle involved in this method is to approximate the integral in (1) by a finite summation, thus reducing the integral equation into a corresponding, finite number of simultaneous differential equations. The utility of the quadrature method therefore depends on the accuracy with which the integral can be represented by a limited summation and by the complexity and number of the resulting differential equations. In so far as actual calculations are concerned the feasibility of the method has been demonstrated only for parallel-plane symmetry.

For parallel-plane symmetry, (1) reduces to

$$\mu \frac{di(Z, \mu)}{dZ} = - N\sigma_t i(Z, \mu) + 2\pi N\sigma_s \int_{-1}^{+1} i(Z, \mu') f_1(\mu, \mu') d\mu', \tag{8}$$

where Z = distance measured in incident direction.

The integral can then be replaced by a quadrature formula such as that of Gauss, yielding a system of n simultaneous linear differential equations:

$$\mu_k \frac{di(Z, \mu_k)}{dZ} = - N\sigma_t i(Z, \mu_k) + 2\pi N\sigma_s \sum_{j=1}^{n} C_j i(Z, \mu_j) f_1(\mu_k, \mu_j) \tag{9}$$

with

$$k = 1, 2 \cdots n.$$

The factors C_j and μ_j are defined by the quadrature formula. The solution of this set of equations gives the intensity, $i(Z, \mu_k)$, *i.e.*, the intensity for any position, Z, and for n discrete directions. For a highly anisotropic distribution function, $f_1(\mu, \mu')$, the number of terms, n, and hence the number of equations must be very large in order to obtain a satisfactory solution.

Quadrature appears to yield nonlinear differential equations for all other geometries.

Spherical Harmonics Method

A standard method of solving the transport equation is by the use of spherical harmonics [8]. The specific intensity $i(\vec{R}, \vec{\Omega})$ is developed as a series of spherical harmonics in $\vec{\Omega}$, such as

$$i(\vec{R}, \vec{\Omega}) = \sum_{m=0}^{\infty} (2m + 1) E_m(\vec{R}) Y_m(\vec{\Omega}) \tag{10}$$

where $Y_m(\vec{\Omega})$ are spherical harmonics of order m and $E_m(\vec{R})$ are functions of the space coordinates $\vec{R}$ only.

Using the orthogonality properties of spherical harmonics, it is possible to reduce the integro-differential equation into a system of simultaneous ordinary differential equations in $E_m(\vec{R})$. For parallel plane symmetry, $Y_m(\vec{\Omega})$ degenerates into the Legendre polynomials, the component equations are linear, and a general solution may be obtained. With spherical symmetry and apparently any other geometry the method yields nonlinear differential equations of such a complicated form that no simple solution is apparent.

An infinite set of algebraic equations can be obtained by taking moments of the space function $E_m(\vec{R})$. Approximate methods have been proposed to obtain a solution from this infinite set [8].

It has been pointed out by Chandrasekhar that the approximation of $i(\vec{R}, \vec{\Omega})$ by a series of spherical harmonics of $2m$ terms is equivalent to the method of quadrature of order $n = 2m$. The difficulties and results of the method of spherical harmonics are therefore the same as those for quadrature.

Reiterative Method

A formal solution of the transport equation may be expressed by a pair of integrals, which can, at least theoretically, yield numerical results by reiteration. The necessary order of reiteration, and the complexity of the integrals, however, limit the usefulness of such a solution.

Briefly, if a "source function" is defined as,

$$q(\vec{R}, \vec{\Omega}) = \int_{\Omega} i(\vec{R}, \vec{\Omega}') f_1(\vec{\Omega}, \vec{\Omega}') d\vec{\Omega}' \tag{11}$$

the transport equation may be reduced to

$$\frac{di(\vec{R}, \vec{\Omega})}{dS_{\Omega}} + N\sigma_t i(\vec{R}, \vec{\Omega}) = N\sigma_s q(\vec{R}, \vec{\Omega}). \tag{12}$$

Direct integration of (12) yields

$$i(\vec{R}, \vec{\Omega}) = N\sigma_s \int q(\vec{R}', \vec{\Omega}) e^{-N\sigma_t |\vec{R}' - \vec{R}|} dS_{\Omega} \tag{13}$$

in which the integration is carried out, along a straight line in the direction $\vec{\Omega}$ from the boundary to the point $\vec{R}$. The pair of equations, (11) and (13) formally define a complete solution of the transport equation.

For numerical calculations, it is more feasible to calculate the contributions to $i(\vec{R}, \vec{\Omega})$ due to the various individual orders of scattering up to the order at which the contribution is negligible and then to take the sum of all orders. Obviously, (11) and (13) can also be interpreted for the individual orders of scattering. In general, the

required integrations cannot be carried out analytically, and the necessary number of reiterations may be very large if the dispersion is dense.

A somewhat simplified form of integral solution may be obtained by introducing an approximate distribution of the kth-scattered radiation estimated from the distribution for single scattering, $f_1(\vec{\Omega}, \vec{\Omega}')$. This approximation, originated by Hartel [9], involves the assumption that the dispersion is so dense that only the radiation scattered from the immediate neighborhood of an obstacle; *i.e.*, having the same distribution as from the obstacle itself, is of importance. The distribution of kth scattered radiation, with respect to direction of the incident radiation, Ω_0, is then

$$f_k(\vec{\Omega}_0, \vec{\Omega}) = \int_{\Omega'} f_{k-1}(\vec{\Omega}_0, \vec{\Omega}') f_1(\vec{\Omega}, \vec{\Omega}') d\Omega'. \tag{14}$$

Using these approximate equations, it is only necessary to evaluate the integrated power due to each order of scattering, and multiply it by the proper distribution function. Even with this approximation, the required calculations are too detailed for practical application.

Approximate Representation of Multiple Scattering by the Classical Diffusion Equations

If the specific intensity, $i(\vec{R}, \vec{\Omega})$, is represented only by a two-term series of spherical harmonics; *i.e.*, by an isotropic function of $\vec{R}$ plus an anisotropic function of $\vec{R}$, such as

$$i(\vec{R}, \vec{\Omega}) = i_0(\vec{R}) + A(\vec{R}) \cdot \vec{\Omega} \tag{15}$$

the transport equation can be degenerated to the classical wave equation, or, in the case of negligible absorption, to the Laplace equation. The solution of these latter equations, together with the appropriate boundary conditions, is relatively simple mathematically, compared to the general solution of the transport equation and has been the subject of considerable study in classical physics. The over-simplified representation of the specific intensity by (22) yields useful results for an infinite dispersion or for several mean free paths from a source or the boundaries of a finite dispersion, but it is inaccurate near a source or boundary where the specific intensity is highly anisotropic.

Specifically, the introduction of (15) permits reduction of (1) to

$$D_f \, div \, grad \, \psi(\vec{R}) + N\sigma_a \psi(\vec{R}) = 0 \tag{16}$$

where the diffusivity

$$D_f = \frac{1}{3N\sigma_t(1 - \omega\mu_m)} \tag{17}$$

with

$$\mu_m = \int_{\Omega'} \mu(\vec{\Omega}, \vec{\Omega}') f(\vec{\Omega}, \vec{\Omega}') d\Omega'. \tag{18}$$

The intensity is obtained from the expression

$$i(\vec{R}, \vec{\Omega}) = \frac{\psi(\vec{R})}{4\pi} - \frac{[grad \, \psi(\vec{R})] \cdot \vec{\Omega}}{4\pi N\sigma_t(1 - \omega\mu_m)}. \tag{19}$$

For no absorption, $\sigma_a = 0$, and (16) reduces to

$$div \, grad \, \psi(\vec{R}) = 0. \tag{20}$$

Complications occur in applying the formal methods of solution of (16) and (20). The general solution is obtained in terms of $\psi(\vec{R})$ or $i(\vec{R}, \vec{\Omega})$, but the boundary conditions in scattering problems are generally specified in terms of the power per unit area crossing a boundary in a given direction and cannot be satisfied by (19). For example, if no power enters a boundary, $i(\vec{R}, \vec{\Omega})$ must be exactly zero for all angles over a hemisphere.

Approximate boundary equations may be derived for certain specific geometries as a substitute for (19). For parallel-plane symmetry, the flux per unit area in the forward direction, E_{Z+}, and the flux per unit area in the backward direction, E_{Z-}, are approximately

$$E_{Z\pm} \cong \frac{\psi(Z)}{4} \mp \frac{1}{6N\sigma_t(1 - \omega\mu_m)} \frac{d\psi(Z)}{dZ} \tag{21}$$

if the third and higher derivations of $\psi(Z)$ are neglected.

It has been suggested that the inaccuracy of the diffusion approximation near a boundary can be mitigated by introducing an imaginary boundary at which $\psi(\vec{R})$ vanishes. For example, from (32), it can be seen that if $E_{Z-} = 0$ at the boundary, $\psi(\vec{R})$ would vanish at a distance beyond the boundary

$$L_t = \frac{2}{3N\sigma_t(1 - \omega\mu_m)} \tag{22}$$

called the linear extrapolation distance. Glasstone and Edlund [5] report that rigorous solution of the transport equation for parallel-plane symmetry yields a coefficient of 0.7104 rather than the 2/3 indicated in (22).

Discrete Flux Representation of Multiple Scattering

An approximate solution of problems in multiple scattering can be obtained if a simplified representation is used for the angular distribution of radiation from a single scattering. If the actual distribution for single scattering is utilized, the transport equation must be solved for all directions at all points in the dispersion. If the angular distribution is represented by one or more discrete components, or by an isotropic component plus one or more discrete components, the transport equation can be reformulated as a set of relatively simple differential equations, and analytical solutions may be possible. Such representations have been proposed by Theissing [11], Fritz [12], and Chu and Churchill [13].

Theissing's Two-Flux Model

Theissing formulated a model for spherical particles in which the scattered radiation is confined to two directions, forward and backward. The continuous distribution and Theissing's model are sketched in Fig. 1(a) and (b). Hartel's approximate distribution for the kth-scattered radiation is utilized, so that the forward and backward components of the distribution function for any order of scattering, k, are given by

$$f_k = 2\pi \int_0^{\pi/2} f_k(\theta) \sin\theta d\theta \tag{23}$$

$$b_k = 2\pi \int_{\pi/2}^{\pi} f_k(\theta) \sin\theta d\theta. \tag{24}$$

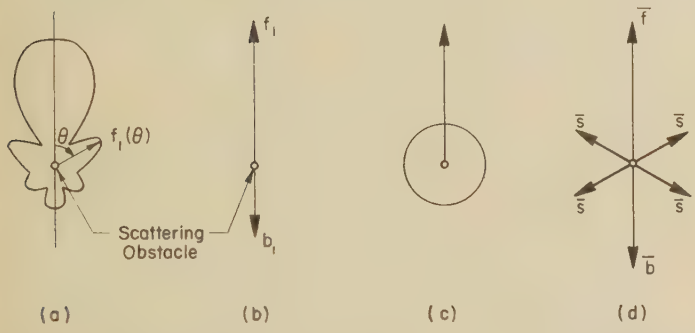


Fig. 1—Angular distribution of radiation from a single scattering. (a) Idealized sketch of a continuous distribution; (b) Theissing's model; (c) Fritz's model; (d) six-flux model.

Energy balances are then written for E_{k+}, the forward power per unit area, and E_{k-}, the backward power per unit area, for each individual order of scattering. If Theissing's analysis is extended to include absorption, the equations for parallel-plane symmetry are:

$$\frac{dE_{0+}}{dZ} + N\sigma_t E_{0+} = 0 \tag{25}$$

$$\frac{dE_{1+}}{dZ} + N\sigma_t E_{1+} = N\sigma_s f_1 E_{0+} \tag{26}$$

$$-\frac{dE_{1-}}{dZ} + N\sigma_t E_1 = N\sigma_s b_1 E_{0+} \tag{27}$$

or in general

$$\frac{dE_{(k+1)+}}{dZ} + N\sigma_t E_{(k+1)+} = N\sigma_s f_{k+1} E_{k+} + N\sigma_s b_{k+1} E_{k-} \tag{28}$$

and

$$-\frac{dE_{(k+1)-}}{dZ} + N\sigma_t E_{(k+1)-} = E\sigma_s f_{k+1} E_{k-} + N\sigma_s b_{k+1} E_{k+} \tag{29}$$

where Z is again the distance through the dispersion.

This set of equations is readily solvable by a stepwise procedure, but the number of orders of scattering which are appreciable, and hence the number of equations that must be solved may be very high for dense dispersions. The accuracy of the method is limited by the simplicity of the model and by the accuracy of Hartel's approximation. The results should be better than those obtained by two-term quadrature, since the weighting factors are evaluated from physical considerations.

Fritz's Two-Stream Model

Fritz represented the radiation from a single scattering by a forward component and an isotropically distributed component as indicated in Fig. 1(c). The medium is divided into layers having a thickness of one-quarter of a mean free path for interception. The diffuse radiation (isotropically distributed) generated in each layer is obtained by numerical integration, as a parallel beam is traced through the medium, layer after layer. An analytical equation is then fitted to a plot of the diffuse radiation generated vs distance. Solutions are then sought for the classical diffusion equation with this empirical expression as a source function. The accuracy of the method should be somewhat better than diffusion theory for which it represents a first order correction.

Six-Flux Model of Chu and Churchill

The representation of the angular distribution for single scattering by six discrete components, one forward, one backward, and four sidewise, as shown in Fig. 1(d), permits reduction of the integro-differential transport equation into a set of six differential equations in terms of the corresponding discrete components of the specific intensity. This set of equations can be solved for many geometries. The solutions are generally detailed, but can be evaluated readily on an automatic computer. The specific intensity can in turn be approximated from the six discrete components.

For randomly-oriented or for symmetrical obstacles, the four sidewise components of the angular distribution function are equal and the forward, backward, and sidewise components can be defined as

$$\bar{f} = 2\pi \int_0^{\pi/2} f_1(\theta) \cos^2\theta \sin\theta d\theta \tag{30}$$

$$\bar{b} = 2\pi \int_{\pi/2}^{\pi} f_1(\theta) \cos^2\theta \sin\theta d\theta \tag{31}$$

and

$$\bar{s} = (1 - \bar{f} - \bar{b})/4. \tag{32}$$

The six discrete components of the specific intensity, $i(\vec{R}, \vec{\Omega})$ are chosen in the direction of and opposite to the three coordinate axes of any orthogonal system of coordinates as shown in Fig. 2. One coordinate is chosen in the direction of the incident wave.

By introduction of $E_{\xi+}$, $E_{\xi-}$, $E_{\zeta+}$, $E_{\zeta-}$, $E_{\eta+}$, and $E_{\eta-}$, the six components of the specific intensity, and $\bar{f}$, $\bar{b}$, and $\bar{s}$, the three components of the angular distribution function into the transport equation, (1), or by writing an energy balance around an elementary volume, a relatively simple set of differential equations are obtained.

Solutions were reported for a parallel-plane wave oblique to a parallel-plane region of scattering and for a point source outside a parallel-plane region. The solutions appear to be more accurate than any of the approximate representations previously considered.

For geometries having a high degree of symmetry such as a parallel-plane wave normally incident upon a parallel-plane region the six-flux model degenerates to a simple two-flux model. This two-flux model is suitable for hand calculations and the solution compares favorably in accuracy with the results of the diffusional representation.

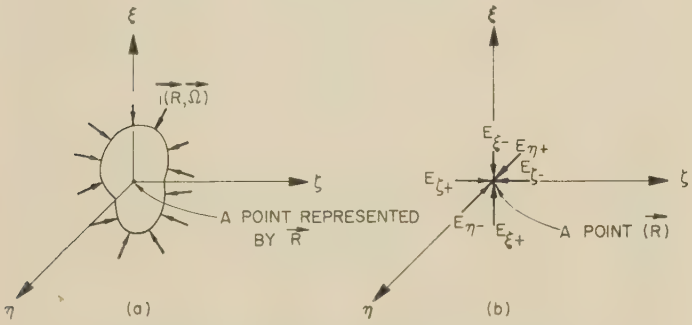


Fig. 2—Representation of specific intensity by discrete components. (a) Actual distribution; (b) discrete representation.

Conclusion

Multiple scattering can be represented in an Eulerian system by an integral equation, called the transport equation, or in a Lagrangian system by a random-walk process. Approximate representation is possible by the classical diffusion equation or by a set of discrete components of the intensity.

Exact solutions for the random-walk formulation can be found only in infinite regions and are difficult to apply for anisotropic distributions. The Monte Carlo technique can be used to obtain particular solutions for any geometry and distribution of scattered power, but for a reasonably accuracy, the required calculations are very extensive, even for a modern automatic computer.

Exact solutions of the transport equation have been found only for isotropic scattering in infinite regions. Approximate solution by quadrature, spherical harmonics, or reiteration can be carried out for parallel-plane boundaries as well as for infinite regions but practical applications are limited to isotropic scattering.

Approximate representation by the diffusion equation permits analytical solution for many geometries and numerical solution for any geometry. This representation is inaccurate near boundaries and for highly anisotropic distributions.

The two-flux model proposed by Chu and Churchill is more accurate than diffusion theory but is practical only for simple geometries. The discrete flux models of Theissing and Fritz are perhaps more accurate but require very extensive calculations. The six-flux model of Chu and Churchill is considerably more accurate and is adaptable to many geometries. However, the analytical solution obtained by the six-flux method is so detailed that calculations are feasible only on an automatic computer. All of the discrete flux methods can be used for anisotropic scattering.

Bibliography

[1] Grosjean, C. C., "The Exact Mathematical Theory of Multiple Scattering of Particles in an Infinite Medium." *Verhandelingen of the Konicklijke Vlaamse Academy* (Belgium), Jaargang XIII, No. 36 (1951).
[2] Grosjean, C. C., "Solution of a Non-isotropic Random Flight Problem in the Case of a Non-isotropic Point Source." *Nuovo Cimento*, Vol. 11 (1954), pp. 11–40.
[3] "Monte Carlo Method." U. S. Government Printing Office, Washington, D. C. (1951), (National Bureau of Standards Applied Mathematics Series-12).
[4] Chandrasekhar, S., "Radiative Transfer." London, Oxford Press, 1950.
[5] Glasstone, S., and Edlund, M. C., "Elements of Nuclear Reactor Theory." New York: Van Nostrand, 1952.
[6] Case, K. M., de Hoffmann, F., Placzek, G., "Introduction to the Theory of Neutron Diffusion." Vol. I U. S. Government Printing Office, Washington, D. C., (1953), (Los Alamos Scientific Laboratory).
[7] Kourganoff, V., "Basic Methods in Transfer Problems." London, Oxford Press, 1952.
[8] Mark, J. C., "The Spherical Harmonic Method, I and II." National Research Council of Canada, CRT-340 and CRT-338 (June, 1947).
[9] Spencer, V., and Fano, U., "Penetration and Diffusion of X-Rays, Calculation of Spatial Distributions by Semi-Asymptotic Methods." *Physical Review*, Vol. 88, (1952), p. 793.
[10] Hartel, W., "Zur Theorie der Lichtstreuung durch trübe Schichten, besonders Trübgläser." *Das Licht*, Vol. 40 (1940), p. 141.
[11] Theissing, H. H., "Macrodistribution of Light Scattered by Dispersions of Spherical Dielectric Particles." *Journal of the Optical Society*, Vol. 40 (1950), p. 232.
[12] Fritz, S., "Scattering and Absorption of Solar Energy by Clouds," Ph.D. Thesis, Massachusetts Institute of Technology, (1953).
[13] Chu, C. M., and Churchill, S. W., "Numerical Solution of Problems in Multiple Scattering of Electromagnetic Radiation." *Journal of Physical Chemistry*, Vol. 59 (1955), p. 855.

Diffraction of Plane Electromagnetic Waves by a Rectangular Aperture*

MICHIO SUZUKI†

Introduction

THE STEADY state problem of the diffraction by an aperture in an infinite plane screen has attracted attention for many years. Available solutions have been restricted to a few cases where the aperture is circular, elliptical or infinitely long and, for rectangular aperture, no adequate theoretical treatment has been given as far as the author knows. This paper is concerned with the theoretical study on the diffraction of plane electromagnetic waves by a rectangular aperture on an infinite plane screen having perfect conductivity. In order to calculate the transmission coefficient the integral-equation method, which can be solved by method of successive approximations, is used when the aperture is long and narrow, on the other hand the variational principle is used when this is not the case.

Integral Equation Formulation for a Rectangular Aperture, and the Transmission Coefficient

We consider an infinite plane conducting screen of infinitesimal thickness, which is perforated by a rectangular aperture τ. A rectangular co-ordinate system is chosen with the origin at the central point of the aperture, and oriented so that the screen lies in x-y plane (Fig. 1).

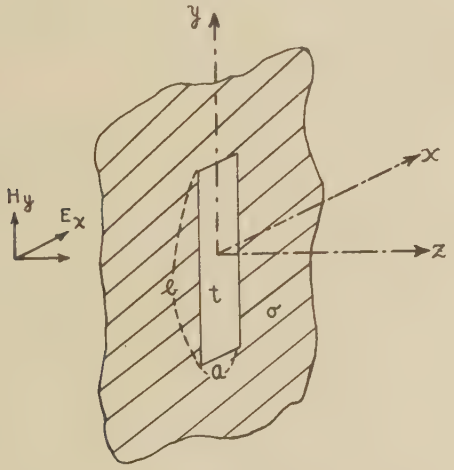


Fig. 1

A plane electromagnetic wave is incident upon the aperture in the half space $Z<0$, and it is desired to investigate the diffracted field. The plane wave $E_x=E_0e^{-jkz}$ (k is free space propagation constant) is incident from the left, and sets up a electric field in the aperture. The incident field is polarized in x-direction only, so that the field in the aperture is also polarized in x-direction only and its amplitude distribution is $\epsilon(x, y)$ (as yet of undetermined form). It is our first object to write the fields in the both regions expressed by $Z\leqq 0$ and $Z\geqq 0$, and then to equate them on the aperture.

For this purpose we introduce a Hertzian vector $\mathbf{\Pi}^*$ of the magnetic type, and express the electromagnetic field **(E, H)** as follows:

$$\mathbf{E} = -j\omega\mu_0\nabla\times\mathbf{\Pi}^*, \qquad \mathbf{H} = \nabla\times\nabla\times\mathbf{\Pi}^* \tag{1}$$

For our problem $\mathbf{\Pi}^*$ is taken to have only a single component Π_y^*. Hence the field components are

$$\begin{aligned} E_x &= j\omega\mu_0\frac{\partial\Pi_y^*}{\partial z} & H_x &= \frac{\partial^2\Pi_y^*}{\partial x\partial y} \\ E_y &= 0 \quad ; & H_y &= \frac{\partial^2\Pi_y^*}{\partial y^2}+k^2\Pi_y^* \\ E_z &= -j\omega\mu_0\frac{\partial\Pi_y^*}{\partial x} & H_z &= \frac{\partial\Pi_y^*}{\partial y\partial z}. \end{aligned} \tag{2}$$

Subscripts 1 and 2 will be used to distinguish the quantities in the regions $Z\leqq 0$ and $Z\geqq 0$ respectively.

For the region $I(Z\leqq 0)$,

$$\Pi_{1y}^* = \Pi_0^* - \Pi_y^*, \tag{3}$$

where

$$\Pi_0^* = \frac{E_0}{\omega\mu_0 k}(e^{-ikz}+e^{ikz}), \tag{4}$$

$$\Pi_y^* = -\frac{1}{2\pi j\omega\mu_0}\int_\tau \epsilon(\xi,\eta)\frac{e^{-jkr}}{r}d\xi d\eta, \tag{5}$$

and for the region $2(Z\geqq 0)$,

$$\Pi_{2y}^* = \Pi_y^*. \tag{6}$$

ξ, η are the coordinates on the aperture, and r is the distance between a point on the aperture and a field point x, y, z;

$$r = [(x-\xi)^2+(y-\eta)^2+(z)^2]^{1/2}.$$

It will be confirmed that

$$\begin{aligned} \lim_{z\to 0}\frac{\partial\Pi_0^*}{\partial z} &= 0 \\ \lim_{z\to 0-}\left(-\frac{\partial\Pi_y^*}{\partial z}\right) &= \begin{cases}\epsilon(x,y) & \tau \\ 0 & \sigma,\end{cases} \\ \lim_{z\to 0+}\left(\frac{\partial\Pi_y^*}{\partial z}\right) &= \begin{cases}\epsilon(x,y) & \tau \\ 0 & \sigma,\end{cases} \end{aligned} \tag{7}$$

and, that Π_{1y}^* and Π_{2y}^* satisfy the wave equation in the respective regions.

* Manuscript received by the PGAP, December 27, 1955.
† College of Engineering, Hokkaido Univ., Sapporo, Japan.

The boundary conditions for the electric field at $Z=0$ are that they should be zero on the metal surface, and continuous and equal to $\epsilon(x, y)$ on the aperture. From (2) and (7) these conditions are satisfied at all. The remaining boundary condition to be satisfied is the continuity of H_y over the aperture. When this has been satisfied, it is easy to see that the boundary conditions for the remaining components are necessarily satisfied.

Equating H_y on each side of the aperture, we find

$$H_{zy}(Z=0) = \frac{\partial^2 \Pi_y^*}{\partial y^2} + k^2 \Pi_y^* = Y_0 E_0 \text{ on } \tau, \quad (8)$$

in which

$$y_0 = \sqrt{\epsilon_0/\mu_0}.$$

Eq. (8) with (5) is an integrodifferential equation for the aperture field distribution $\epsilon(x, y)$.

Now let us introduce the transmission coefficient of the aperture. The transmission coefficient T of the aperture is defined as the ratio of the transmitted energy flux W through the aperture to the incident energy flux w upon the aperture, and expressed as follows:

$$T = \frac{W}{w}, \quad (9)$$

in which

$$W = \frac{1}{2} \operatorname{Re} \int_\tau \epsilon(x, y) \tilde{H}_{zy}(x, y, 0) dx dy, \quad (10)$$

$$w = \frac{1}{2} ab y_0 E_0^2, \quad (11)$$

where $\tilde{H}_{2y}$ is the conjugate complex of H_{2y} and a and b are the sizes of the aperture as shown in Fig. 1.

Let us eliminate $\tilde{H}_{2y}$ in (10) by means of (8), then (9) becomes

$$T = \frac{1}{abE_0} \operatorname{Re} \int_\tau \epsilon(x, y) dx dy. \quad (12)$$

The Case When the Aperture Is Long and Narrow

At first we consider the case when the slot is long in y-direction and narrow in x-direction. The solution of the partial differential (8) will be the form

$$\Pi_y^*(x, y) = \frac{y_0}{k^2} [E_0 + C(x) e^{-iky} + D(x) e^{iky}] \quad (13)$$

where $C(x)$ and $D(x)$ are some functions of x.

If the width a of the slot is small compared with wavelength and the primary wave is incident perpendicular to the aperture, it will be confirmed that we can put $C(x)=C(0)$ $(=C)$, $D(X)=D(0)$ $(=D)$ and $C=D=B/2$ (B is the unknown constant). Equating the right-hand sides of (13) and (5), we find

$$\int \epsilon(\xi, \eta) \frac{e^{-jkr}}{r} d\xi d\eta = -2\pi j \omega \mu_0 \frac{y_0}{k^2} [E_0 + B \cos ky]. \quad (14)$$

This is the integral equation for $\epsilon(x, y)$. After finding of $\epsilon(x, y)$, T will be determined by (12). Referring to Gray's paper [1] let us rewrite the left-hand side of (14) as follows:

$$\int_\tau \epsilon(\xi, y) \frac{\cos kr}{r} d\xi d\eta + \int_\tau \frac{\{\epsilon(\xi, \eta) - \epsilon(\xi, y)\} \cos kr - j\epsilon(\xi, \eta) \sin kr}{r} d\xi d\eta$$

and $r=\sqrt{(x-\xi)^2+(y-\eta)^2}$ in the second integral may be written as $|y-\eta|$ because b is much greater than a. Using these results, (14) becomes

$$\int_{-a/2}^{a/2} \epsilon(\xi, y) K(x, \xi; y) d\xi + \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} \frac{\{\epsilon(\xi, \eta) - \epsilon(\xi, y)\} \cos k|y-\eta| - j\epsilon(\xi, \eta) \sin k|y-\eta|}{|y-\eta|} d\xi d\eta = -\frac{2\pi i}{k} (E_0 + B \cos ky), \quad (15)$$

where

$$K(x, \xi; y) = \int_{-b/2}^{b/2} \frac{\cos kr}{r} \alpha\eta = \log \frac{\sqrt{\left(\frac{b}{2}+y\right)^2 + (x-\xi)^2} + \left(\frac{b}{2}+y\right)}{\sqrt{\left(\frac{b}{2}-y\right)^2 + (x-\xi)^2} - \left(\frac{b}{2}-y\right)} - C_{in}\left\{k\left(\frac{b}{2}+y\right)\right\} - C_{in}\left\{k\left(\frac{b}{2}-y\right)\right\},$$

$$C_{in}(x) = \log x + 0.5772 - C_i(x),$$

$C_i(x)$; integral cosine.

Since the function $K(x, \xi; y)$ is slowly varying function of y, it may be replaced by its mean value

$$K(x,\xi) = \frac{1}{b}\int_{-b/2}^{b/2} K(x,\xi;y)dy$$
$$= 2\left[\log\frac{b}{x-\xi} + \log x - C_{in}(2\beta) - \frac{\sin 2\beta}{2\beta}\right], \quad (16)$$

where

$$\beta = \frac{\pi b}{\lambda}.$$

On changing the variables x and ξ to θ and θ', respectively, by putting $x = a/2\cos\theta$ and $\xi = a/2\cos\theta'$, where θ and θ' vary from 0 to π, and using the formula

$$\sum_1^\infty \frac{\cos m\theta \cos m\theta'}{m} = -\frac{1}{2}\log 2\left|\cos\theta - \cos\theta'\right|,$$

we get

$$K(x,\xi) = \Omega_G + 4\sum_1^\infty \frac{\cos m\theta \cos m\theta'}{m}, \quad (17a)$$

where

$$\Omega_G = 2\left[\log\frac{4b}{a} - C_{in}(2\beta) - \frac{\sin 2\beta}{2\beta} + \log 2\right]. \quad (7b)$$

Expanding s in $\epsilon(\xi, y)$ in a Fourier series as follows:

$$\sin\theta'\epsilon(\xi,y) = \sum_0^\infty C_m(y)\cos m\theta' \quad (18)$$

and substituting this into (16) and making some simplifications, we get

$$-\frac{\pi a}{2}\left[\Omega_G C_0(y) + 2\sum_1^\infty \frac{C_m(y)\cos m\theta}{m}\right]$$
$$= \frac{2\pi j}{k}(E_0 + B\cos ky) + \frac{\pi a}{2}\int_{-b/2}^{b/2} \frac{\{C_0(\eta) - C_0(y)\}\cos k|y-\eta| - jC_0(\eta)\sin k|y-\eta|}{|y-\eta|}d\eta. \quad (19)$$

Equating the coefficients of $\cos m\theta (m = 0, i, \cdots)$ in both sides of (19), we get the following relations:

$$C(y) = C(y)\cdots = 0,$$
$$C_0(y) = \frac{f}{\Omega_G}(E_0 + B\cos ky) - \frac{1}{\Omega_G}\int_{-b/2}^{b/2} \frac{\{C_0(\eta) - C_0(y)\}\cos k|y-\eta| - jC_0(\eta)\sin k|y-\eta|}{|y-\eta|}d\eta. \quad (20)$$

where

$$f = -j\frac{4}{ka}.$$

Eq. (20) is also the integral equation for $C_0(y)$, and can be solved by means of the method of successive approximations which Hállen used in his paper concerning the theory of electric current antenna. Thus we get the following solution of $C_0(y)$.

$$C_0(y) = \frac{f}{\Omega_G}\left[E_0 + B\cos ky + \frac{1}{\Omega_G}\left\{jE_0L(y) - \frac{B}{2}M(y) + j\frac{B}{2}N(y)\right\}\right], \quad (21)$$

where

$$L(y) = S_i\left\{k\left(\frac{b}{2}+y\right)\right\} + S_i\left\{k\left(\frac{b}{2}-y\right)\right\},$$
$$M(y) = \sin ky\left[S_i\left\{2k\left(\frac{b}{2}+y\right)\right\} - S_i\left\{2k\left(\frac{b}{2}-y\right)\right\}\right]$$
$$- \cos ky\left[C_{in}\left\{2k\left(\frac{b}{2}-y\right)\right\} + C_{in}\left\{2k\left(\frac{b}{2}-y\right)\right\}\right],$$
$$N(y) = \cos ky\left[S_i\left\{2k\left(\frac{b}{2}+y\right)\right\} + S_i\left\{2k\left(\frac{b}{2}-y\right)\right\}\right]$$
$$+ \sin ky\left[C_{in}\left\{2k\left(\frac{b}{2}+y\right)\right\} - C_{in}\left\{2k\left(\frac{b}{2}-y\right)\right\}\right],$$

$S_i(x)$; integral—sine.

Now the boundary condition for $\epsilon(x, y)$ at $y = \pm b/2$ is $\epsilon(x, y = \pm b/2) = 0$; that is, $C_0(\pm b/2) = 0$.

From this condition the coefficient B can be determined. Hence the electric field distribution $\epsilon(x, y)$ in the aperture becomes as follows:

$$\epsilon(x,y) = \frac{C_0(y)}{\sin\theta} = \frac{C_0(y)}{\sqrt{1-\cos^2\theta}} = \frac{C_0(y)}{\sqrt{1-\left(\frac{2X}{a}\right)^2}}$$
$$= -\frac{4jE_0}{ka\Omega_G}\frac{1}{\sqrt{1-\left(\frac{2X}{a}\right)^2}}\left[1 - \delta\cos ky + \frac{1}{\Omega_G}\left\{jL(y) + \frac{\delta}{2}M(y) - j\frac{\delta}{2}N(y)\right\}\right], \quad (22)$$

where

$$\delta = \frac{1 + \frac{j}{\Omega_G} S_i(2\beta)}{\cos\beta - \frac{1}{2\Omega_G}\left[\sin\beta S_i(4\beta) - \cos\beta C_{in}(4\beta) - j\left\{\cos\beta S_i(4\beta) + \sin\beta C_{in}(4\beta)\right\}\right]}. \tag{23}$$

The aperture electric field distribution in x-direction is given by $[1-(2\alpha/a)e]^{-1/2}$, and this distribution seems to be correct, since the width of the aperture is so small compared with wavelength that the field distribution in the same direction may be considered just the same as static field.

Substituting the $\epsilon(x, y)$ from (22) in (12) and carrying out the integration, the transmission coefficient T is given as follows:

$$T = \frac{\pi}{\alpha\Omega_G}\operatorname{Im}\left[1 - \delta\frac{\sin\beta}{\beta} + \frac{1}{\Omega_G}\left\{js_0 - \frac{\delta}{4\beta}e^{j\beta}(q_1 + jq_2)\right\}\right], \tag{24}$$

where

$$s_0 = 2S_i(2\beta) - \frac{I - \cos 2\beta}{\beta},$$

$$q_1 = 2\{\sin 2\beta C_{in}(4\beta) - \cos 2\beta S(4\beta) + S_i(6\beta) - S_i(2\beta)\},$$

$$q_2 = 2\{\sin 2\beta S_i(4\beta) + \cos(2\beta)C_{in}(4\beta) + C_{in}(2\beta) - C_{in}(6\beta)\},$$

and Ω_G and δ are given in (17a) and (23) respectively.

Now if we rewrite (14) as follows:

$$\int_\tau \frac{\epsilon(\xi, y)}{r}d\xi\alpha\eta + \int_\tau \frac{\epsilon(\xi, \eta)e^{-jkr} - \epsilon(\xi, y)}{r}d\xi d\eta = -\frac{2\pi j}{k}(E_0 + B\cos ky),$$

another solution for T is obtained by the similar way.

This is as follows:

$$T = \frac{\pi}{\alpha\Omega_H}\operatorname{Im}\left[1 - \delta\frac{\sin\beta}{\beta} + \frac{1}{\Omega_H}\left\{2E(2\beta) - 2 + \frac{2(1 - e^{-j2\beta}) - \delta F}{j2\beta}\right\}\right],$$

where

$$F = e^{+j\beta}E(4\beta) - j2e^{-j\beta}S_i(2\beta),$$

$$E(x) = C_{in}(x) + jS_i(x),$$

$$\Omega_H = 2\log\frac{4b}{a}$$

$$\delta = \frac{1 + \frac{1}{\Omega_H}E(2\beta)}{\cos\beta + \frac{e^{j\beta}}{2\Omega_H}E(4\beta)}. \tag{25}$$

These formulas are available for the case when the aperture is long and narrow.

The Solutions by the Variational Method

In this section we will find the solution by using the variational method. Using (5), (8) and (12), the stationary expression for the transmission coefficient T of the aperture, as derived by Levine and Schwinger [2], is

$$T = \frac{(2\pi)^2}{\lambda ab}\operatorname{Im}\left(\frac{G^2}{F}\right),$$

$$G = \int_\tau \epsilon(x, y)dxdy,$$

$$F = \int_\tau\int_\tau \epsilon(x, y)\epsilon(\xi, \eta)\left(\frac{\partial^2}{\partial y^2} + k^2\right)\frac{e^{-jkr}}{r}d\xi d\eta dxdy, \tag{26}$$

where the notation Im means the imaginary part of. F is the quadruple integral and this is very troublesome to evaluate, but it can be treated as follows. Now let us introduce the integral representation for e^{-jkr}/r, that is

$$\frac{e^{-jk\sqrt{(x-\xi)^2+(y-\eta)^2}}}{\sqrt{(x-\xi)^2+(y-\eta)^2}} = \frac{1}{2\pi}\int\int_{-\infty}^{\infty}\frac{e^{-j\alpha(y-\eta)-\sqrt{\alpha^2+\beta^2-k^2}|x-\xi|}}{\sqrt{\alpha^2+\beta^2-k^2}}d\alpha d\beta = \frac{2}{\pi}\int\int_0^{\infty}\frac{\cos\alpha(y-\eta)e^{-\sqrt{\alpha^2+\beta^2-k^2}|x-\xi|}}{\sqrt{\alpha^2+\beta^2-k^2}}d\alpha d\beta. \tag{27}$$

As y and η vary from $-b/2$ to $b/2$, $z = y-\eta$ varies from $-b$ to b. Expanding $\cos\alpha Z$ in a Fourier series in the range from $-b$ to b as follows:

$$\cos\alpha Z = \sum_{n=0}^{\infty} A_n\cos q_n Z, \tag{28}$$

where

$$A_0 = \frac{1}{2b}\int_0^b \cos\alpha Z dZ = \frac{\sin\alpha b}{\alpha b},$$

$$A_n = \frac{1}{b}\int_0^b \cos\alpha Z\cos q_n Z dZ = \frac{2\sin\alpha b}{b}\frac{(-1)^n\alpha}{\alpha^2 - q_n^2},$$

$$q_n = \frac{n\pi}{b},$$

and inserting this into (27), we get

$$\frac{e^{-jkr}}{r} = \frac{2}{\pi}\frac{2}{b}\sum_0^{\infty}(-1)^n\epsilon_n\cos q_n(y - \hat{\eta})\int_0^{\infty} I_n(\beta)d\beta, \tag{29}$$

where

$$\epsilon_n = \tfrac{1}{2}(n = 0), \qquad 1(n \neq 0)$$

$$I_n(\beta) = \int_0^\infty \frac{\alpha \sin \alpha\beta e^{-\sqrt{\alpha^2+\beta^2-k^2}|x-\xi|}}{(\alpha^2 - q_n^2)\sqrt{\alpha^2+\beta^2-k^2}}\, d\alpha. \tag{30}$$

Integrating the following functions $f_+(\mu)$ and $f_-(\mu)$

$$f_\pm = \frac{\mu e^{\pm j b\mu} e^{-\sqrt{\mu^2+\beta^2-k^2}|x-\beta|}}{(\mu^2 - q_n^2)\sqrt{\mu^2+\beta^2-k^2}}$$

along the contours C_1 and C_2 respectively as shown in Fig. 2, it will be shown that [3]

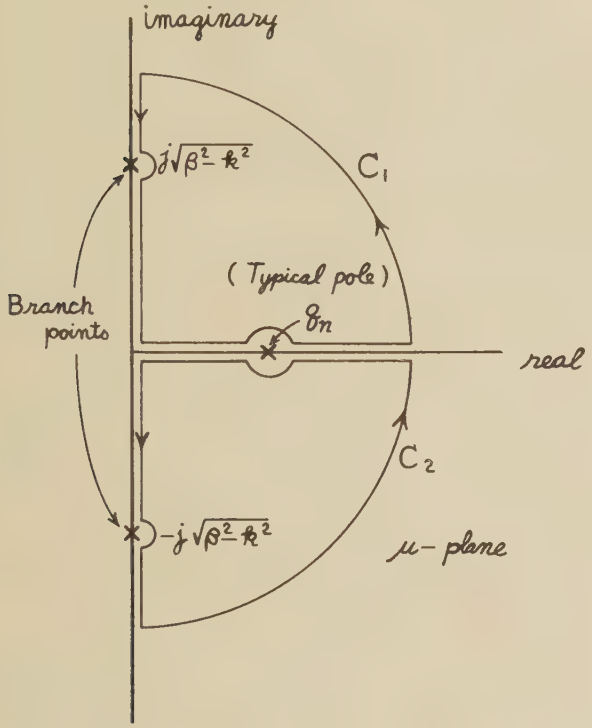


Fig. 2

$$I_n(\beta) = (-1)^n \frac{\pi}{2} \frac{e^{-\sqrt{\beta^2-\alpha_n^2}|x-\xi|}}{\sqrt{\beta^2-\alpha_n^2}}, \tag{31}$$

where

$$\alpha_n^2 = k^2 - q_n^2.$$

Inserting (31) into (29), we get

$$\frac{e^{-jkr}}{r} = \frac{2}{b}\sum_0^\infty \epsilon_n X_n(|x-\xi|)\cos q_n(y-n), \tag{32}$$

where

$$X_n(|x-\xi|) = \int_0^\infty \frac{e^{-\sqrt{\beta^2-\alpha_n^2}|x-\xi|}}{\sqrt{\beta^2-\alpha_n^2}}\, d\beta. \tag{33}$$

The integral $X_n(|x-\xi|)$ can be evaluated as follows:

1) If $\alpha_n^2<0$, putting $\alpha_n^2 = -\alpha_n'^2 (\alpha_n'^2>0)$ and changing the variable from β to t by putting $\sqrt{\beta^2+\alpha_n'^2} = \alpha_n' t$, we get

$$X_n(|x-\xi|) = \int_1^\infty \frac{e^{-\alpha_n'|x-\xi|s}}{\sqrt{t^2-1}}\, dt = K_0(\alpha_n'|x-\xi|) \tag{34}$$

by using the well-known formula, where $K_0(x)$ denotes the modified Bessel function of the order zero.

2) If $\alpha_n^2>0$, the integral (33) can be divided into two integrals as follows:

$$X_n(|x-\xi|) = -j\int_0^{\alpha_n} \frac{e^{-j\sqrt{\alpha_n^2-\beta^2}|x-\xi|}}{\sqrt{\alpha_n^2-\beta^2}}\, d\beta + \int_{\alpha_n}^\infty \frac{e^{-\sqrt{\beta^2-\alpha_n^2}|x-\xi|}}{\sqrt{\beta^2-\alpha_n^2}}\, d\beta..$$

Changing the variable from β in the first integral to ϕ by putting $\sqrt{\alpha_n^2-\beta^2} = \alpha_n \cos\phi$, the first integral becomes

$$-j\int_0^{\pi/2} e^{-j\alpha_n|x-\xi|\cos\phi} d\phi. \tag{35}$$

Employing the Jacobi's expansion in the series of Bessel coefficients,

$$e^{-jZ\cos\phi} = J_0(Z) + 2\sum_{m=1}^\infty (-1)^m J_{2m}(Z)\cos 2m\phi - j2\sum_{m=0}^\infty (-1)^m J_{2m+1}(Z)\cos(2m+1)\phi \tag{36}$$

and inserting this expansion into (35) and integrating term by term, we get

$$\frac{\pi}{2}J_0(\alpha_n|x-\xi|) - 2j\sum_{m=0}^\infty (-1)^m \frac{J_{2m+1}(\alpha_n|x-\xi|)}{2m+1}.$$

Thus $X_n(|x-\xi|)$ becomes as follows:

$$X_n(|x-\xi|) = -j\frac{\pi}{2}J_0(\alpha_n|x-\xi|) - 2\sum_{m=0}^\infty -(1)^n \frac{J_{2m+1}(\alpha_n|x-\xi|)}{2m+1} + K_0(\alpha_n|x-\xi|). \tag{37a}$$

If $\alpha = \pi a/\lambda$ is very small compared to unity, $X_n(|x-\xi|)$ may be approximated as follows:

$$X_n(|x-\xi|) \doteqdot \delta_n\left[-j\frac{\pi}{2}\left\{1-\frac{1}{4}(x-\xi)^2\right\} - \alpha_n|x-\xi|\right] + \log\frac{1\cdot 12}{|(x-\xi)\alpha_n|}, \tag{37b}$$

where $\delta_n = 1$, if α_n is real, and $\delta_n = 0$ if Ω_n is imaginary.

Now let us assume that $\epsilon(x, y)$ will be a product of the functions of x and y separately, that is,

$$\epsilon(x, y) = \epsilon_1(x)\epsilon_2(y). \tag{38}$$

Using (32) and (38), (26) becomes, after some simplifications, as follows:

$$T = \frac{\pi}{2\alpha} \operatorname{Im}\left(\frac{1}{D}\right), \tag{39}$$

where

$$D = \sum_0^\infty \epsilon_n \left(\frac{\alpha_n}{k}\right)^2 \Gamma_n \Delta_n, \tag{40}$$

$$\Gamma_n = \frac{\int\int_{-a/2}^{a/2} \epsilon_1(x)\epsilon_1(\xi) X_n(|x-\xi|)dxd\xi}{\left[\int_{-a/2}^{a/2} \epsilon_1(x)dx\right]^2}, \tag{41}$$

$$\Delta_n = \frac{\int\int_{-b/2}^{b/2} \epsilon_2(y)\epsilon_2(\eta)\cos\alpha_n(y-\eta)dyd\eta}{\left[\int_{-b/2}^{b/2} \epsilon_2(y)dy\right]^2}. \tag{42}$$

1) At first let us consider the case when $\alpha \ll 1$.

In this case we will take as a trial distribution of $\epsilon_1(x)$ the following function:

$$\epsilon_1(x) = \left[1 - \left(\frac{2x}{a}\right)^2\right]^{-1/2}. \tag{43}$$

Inserting (43) with (37a) into (40) and carrying out the integrations, we get

$$\Gamma_n = \delta_n\left[-j\frac{\pi}{2}\left\{1 - \frac{1}{4}\left(\frac{\alpha_n a}{2}\right)^2\right\} - \left(\frac{2}{\pi}\right)^2\left(\frac{\alpha_n a}{2}\right)\right] + \log\frac{2\cdot 24}{|\alpha_n|\, a/2}. \tag{44}$$

2) If we put $\epsilon_1(x) = I$, (40) is evaluated as follows:

Let $F(\lambda)$ be a function of λ and $\int_0^\lambda F(\mu)d\mu = F_1(\lambda)$, then we find, by integration by parts,

$$\int_0^D \int_0^\lambda F(\mu)d\mu d\lambda = \int_0^D F_1(\lambda)d\lambda = [\lambda F_1(\lambda)]_0^D - \int_0^D \lambda F(\lambda)d\lambda = \int_0^D (D-\lambda)F(\lambda)d\lambda.$$

Similarly,

$$\int_0^D \int_0^{D-\lambda} F(\mu)d\lambda d\mu = \int_0^D (D-\lambda)F(\lambda)\alpha(\lambda)\theta.$$

Using these results to carry out the x- and ξ-integrations in (41), we get

$$\Gamma_n = \frac{2}{a^2}\left[\sigma_n\left\{-j\frac{\pi}{2}\int_0^a (a-\lambda)J_0(\alpha_n\lambda)d\lambda - 2\sum_{m=0}^\infty \frac{1}{2m+1}\int_0^a (a-\lambda)J_{2m+1}(\alpha_n\lambda)d\lambda\right\} + \int_0^a (a-\lambda)K_0(\alpha_n\lambda)d\lambda\right].$$

Using the following formulas:

$$\frac{d}{dx}\left[x^\rho J_\rho(x)\right] = x^\rho J_{\rho-1}(x),$$

$$\frac{d}{dx}\left[x^\rho K_\rho(x)\right] = -x^\rho K_{\rho-1}(x),$$

$$\int J_\rho(x)dx = 2\sum_{\nu=0}^\infty J_{\rho+2\nu+1}(x),$$

and employing the power series solutions for $Jp(x)$ and the ascending series solution for $K_0(x)$, we get

$$\Gamma_n = \frac{2}{\alpha_n a}\left[\delta_n\left\{-j\frac{\pi}{2}\left(2\sum_{\nu=0}^\infty J_{2\nu+1}(\alpha_n a) - J_1(\alpha_n a)\right) - 2\sum_{m=0}^\infty \frac{1}{2m+1}\left(2\sum_{\nu=0}^\infty J_{2m+2\nu+2}(\alpha_n a) + P_{2m+1}(\alpha_n a)\right)\right\} + Q(\alpha_n a) + \frac{\alpha_n a K_1(\alpha_n a) - 1}{(\alpha_n a)^2}\right], \tag{45}$$

where

$$P_{2m+1}(\alpha_n a) = (\alpha_n a)\sum_{\nu=0}^\infty \frac{\left(\frac{\alpha_n a}{2}\right)^{2m+2\nu+1}}{\nu!(2m+\nu+1)!(2m+2\nu+3)},$$

$$Q(\alpha_n a) = -\log\left(\frac{\alpha_n a}{2}\right)\sum_{m=0}^\infty \frac{\left(\frac{\alpha_n a}{2}\right)^{2m}}{(m!)^2(2m+1)} + \sum_{m=0}^\infty \frac{\left(\frac{\alpha_n a}{2}\right)^{2m}}{\{m!(2m+1)\}^2} + \sum_{m=0}^\infty \frac{\psi(m)}{(m!)^2(2m+1)}\left(\frac{\alpha_n a}{2}\right)^{2m}$$

$$\psi(m) = -0.5772 + I + \frac{1}{2} + \cdots + \frac{1}{m}.$$

3) If we put

$$\epsilon_2(y) = \left[1 - \left(\frac{2y}{b}\right)^2\right]^{1/2},$$

(42) becomes

$$\Delta_n = \left(\frac{2}{\pi}\right)^2 \int\int_0^\pi \cos\left\{\frac{n\pi}{2}(\cos\theta - \cos\theta')\right\} \cdot \sin^2\theta \sin^2\theta' d\theta d\theta'$$

$$= \left[J_0\left(\frac{n\pi}{2}\right) + J^2\left(\frac{n\pi}{2}\right)\right]^2 = 4\left\{\frac{J_1\left(\frac{n\pi}{2}\right)}{\frac{n\pi}{2}}\right\}^2. \tag{46}$$

after changing the variables from y and η to θ and θ' respectively by putting $y=(b/2)\cos\theta$ and $\eta=(b/2)\cos\theta'$ and using the Jacobi's expansion in series of Bessel coefficients.

4) If we put $\epsilon_2(y)=\cos\beta-\cos ky$, then (42) becomes as follows:

$$\Delta_0 = 1$$

$$\Delta_{q_n=k} = \frac{(\beta/2)^2}{(\beta\cos\beta-\sin\beta)^2},$$

$$\Delta_{2n+1} = \left\{\frac{\frac{2}{\pi}\beta\cos\beta}{\beta\cos\beta+\sin\beta}\right\}^2 \cdot\left[\frac{1}{(2n+1)\left\{1-\left(\frac{2n+1\pi}{2\beta}\right)^2\right\}}\right],$$

$$\Delta_{2n} = \left(\frac{\sin\beta}{\beta\cos\beta-\sin\beta}\right)^2 \frac{1}{\left\{1-\left(\frac{2n\pi}{2\beta}\right)^2\right\}^2}. \tag{47}$$

The results so far obtained may be summarized as follows:

$$T = \frac{\pi}{2\alpha}\,\mathrm{Im}\left(\frac{1}{D}\right)$$

where

$$D = \frac{1}{2}\Delta_0\Gamma_0 + \sum_1^\infty\left\{1-\left|\frac{\frac{n\pi}{2}}{\beta}\right|^2\right\}\Gamma_n\Delta_n,$$

with

$$\Gamma_n = \delta_n\left[-j\frac{\pi}{2}\left\{1-\frac{1}{4}\left(\frac{\alpha_n a}{2}\right)^2\right\}-\left(\frac{2}{\pi}\right)^2\left(\frac{\alpha_n a}{2}\right)\right] + \log\frac{2\cdot 24}{|\alpha_n|\,a/2}, \tag{48b}$$

$$\Gamma_n = \frac{2}{\alpha_n a}\left[\delta_n\left\{-j\frac{\pi}{2}\left(2\sum_{\nu=0}^\infty J_{2\nu+1}(\alpha_n a)-J_1(\alpha_n a)\right) - 2\sum_{m=0}^\infty\frac{1}{2m+1}\left(2\sum_{\nu=0}^\infty J_{2m+2\nu+2}(\alpha_n a)-P_{2m+1}(\alpha_n a)\right)\right\} + Q(\alpha_n a)+\frac{\alpha_n a K_1(\alpha_n a)-1}{(\alpha_n a)^2}\right],$$

$$P_{2m+1}(\alpha_n a) = (\alpha_n a)\sum_{\nu=0}^\infty\frac{\left(\frac{\alpha_n a}{2}\right)^{2m+2\nu+1}}{\nu!(2m+\nu+1)!(2m+2\nu+3)},$$

$$Q(\alpha_n a) = -\log\left(\frac{\alpha_n a}{2}\right)\sum_{m=0}^\infty\frac{\left(\frac{\alpha_n a}{2}\right)^{2m}}{(m!)^2(2m+1)} + \sum_{m=0}^\infty\frac{\left(\frac{\alpha_n a}{2}\right)^{2m}}{\{m!(2m+1)\}^2} + \sum_{m=0}^\infty\frac{\psi(m)\left(\frac{\alpha_n a}{2}\right)^{2m}}{(m!)^2(2m+1)}$$

$$\psi(m) = -0.5772+1+\frac{1}{2}+\cdots+\frac{1}{m},$$

$$\Delta_n = 4\left\{\frac{J_1\left(\frac{n\pi}{2}\right)}{\frac{n\pi}{2}}\right\}^2, \tag{48c}$$

$$\Delta_0 = 1,$$

$$\Delta_{q_n=k} = \frac{(\beta/2)^2}{(\beta\cos\beta-\sin\beta)^2},$$

$$\Delta_{2n+1} = \left\{\frac{\frac{2}{\pi}\beta\cos\beta}{\beta\cos\beta-\sin\beta}\right\}^2 \cdot\left[\frac{1}{(2n+1)\left\{1-\left(\frac{2n+1\pi}{2\beta}\right)^2\right\}^2}\right]^2,$$

$$\Delta_{2n} = \left(\frac{\sin\beta}{\beta\cos\beta-\sin\beta}\right)^2 \frac{1}{\left\{1-\left(\frac{2n\pi}{2\beta}\right)^2\right\}^2}. \tag{48d}$$

5) Finally, if we take $\epsilon_2(y)=\cos q_1 y+A\cos q_3 y$ and determine the constant A by putting $\partial T/\partial A=0$, then (39) becomes as follows, after some simplifications,

$$T = \frac{\pi}{2\alpha}\,\mathrm{Im}\left[\frac{t_{11}+t_{33}-2t_{13}}{t_{11}t_{33}-t_{13}^2}\right], \tag{49}$$

where

$$t_{\nu\nu}' = \frac{1}{2}\Gamma_0 + \sum_1^\infty\left\{1-\left|\frac{\frac{n\pi}{2}}{\beta}\right|^2\right\}\Gamma_n\Delta_{n;\nu,\nu}',$$

with

$$\Delta_{1;1,1} = \frac{\pi^2}{16}, \qquad \Delta_{1;1,3} = \Delta_{3;1,3} = 0,$$

$$\Delta_{3;3,3} = 9\left(\frac{\pi^2}{16}\right),$$

$$\Delta_{n;1,1} = \frac{1}{(n^2-1)^2} \qquad \Delta_{n;1,3} = \frac{9}{(n^2-1)(n^2-3)},$$

$$\Delta_{n;3,3} = 9\left(\frac{3}{n^2-3^2}\right)^2,$$

and Γ_n is given in (48a) or (48b).

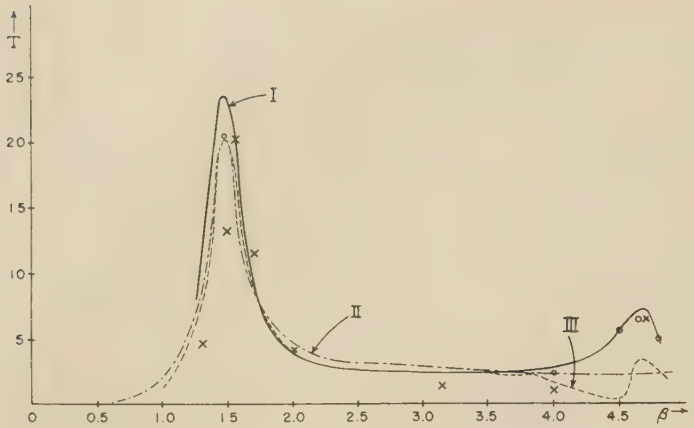


Fig. 3— $\alpha=0.1$

I ——— Successive approximation due to Gray

× Successive approximation due to Hállen

II — - — - — Variational approximation, based on aperture field of the form

$$\epsilon(x, y) = \left[1-\left(\frac{2x}{a}\right)^2\right]^{-1/2}\left[1-\left(\frac{2y}{b}\right)^2\right]^{1/2}$$

III - - - - - Variational approximation, based on aperture field of the form

$$\epsilon(x, y) = \left[1-\left(\frac{2x}{a}\right)^2\right]^{-1/2}[\cos\beta - \cos ky]$$

○ Variational approximation, based on aperture field of the form

$$\epsilon(x, y) = \left[1-\left(\frac{2x}{a}\right)^2\right]^{-1/2}\left[\cos\frac{\pi}{b}y + C\cos\frac{3\pi}{b}y\right].$$

Numerical Examples

In this section two numerical examples are shown in Fig. 3 and Fig. 4. The transmission coefficient is plotted in Fig. 3 as a function of β in the case $\alpha=0.1$. The curve I is calculated from (24) which was obtained by means of the integral equation method.

The curve II and III are calculated from (48a) with (48c), and (48) with (48d) respectively. The points indicated by crosses are obtained by (25), and the points indicated by circles are obtained by (49) with (26).

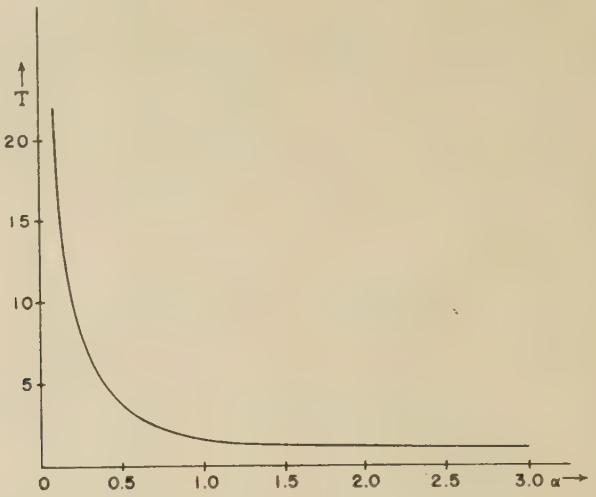


Fig. 4—$\beta=1.5$. Variational approximation based on aperture field of the form $\epsilon(x, y)=(\cos\beta-\cos ky)$.

When the length b of the slot is approximately equal to $\lambda/2$, $3/2\lambda$, $5/2\lambda$, $\cdots$, the resonance phenomena occur. At these points the transmission coefficient takes the largest values.

The transmission coefficient is plotted in Fig. 4 as a function of α in the case $\beta=1.5$. This curve which is obtained from (48b) with (48d) approaches to the value unity as it is increased to infinity.

Acknowledgement

The author wishes to express his sincere thanks to Professor T. Matsumoto for his encouragement.

Bibliography

[1] Gray, M. C., *Journal of Applied Physics*, Vol. 15 (1944), p. 61.
[2] Schwinger, J., *Physical Review*, Vol. 74 (1948), p. 958.
[3] Moriguchi, N., *The Journal of the Institute of Electrical Commercial Engineering of Japan*, Vol. 38 (1955), p. 552.
[4] Watson, The Theory of Bessel Functions.
[5] Tai, C. T., *Journal of Applied Physics*, Vol. 23 (1952), p. 909.

The Amplitude Concept of an Electromagnetic Wave and Its Application to Junction Problems in Waveguides*

JEAN A. ORTUSI†

IT IS generally known that the concept of complex amplitude of electric and magnetic fields is widely used in mathematical physics.

In this paper, it will be shown that this concept can be extended to guided waves and that it enables various functional properties of junctions used in waveguides to be computed and represented.

* Manuscript received by the PGAP, June 10, 1955; revised manuscript received December 12, 1955.

† Compagnie Generale de T.S.F., Paris, France.

Calculation of a Guided Wave

Denoting by $U(x, y)$, the complex amplitude of the longitudinal component of the hertzian vector, it can be readily shown that the longitudinal and transverse electric and magnetic fields are proportional to $U(x, y)$ or $\nabla U_{(x,y)}$.

A guided wave will be defined as *normalized* when the total flux of the Poynting vector in direction Oz, integrated during a period of time, is equal to unit power. The concept of a normalized guided wave allows for the unambiguous definition of the complex amplitude of a guided wave. A nonnormalized wave of a given type is defined by the value of the function U:

$$U = CV(x, y). \tag{1}$$

C is a constant and $V(x, y)$ is obtained by solving the propagation equation and taking into account the boundary conditions:

$$V = 0$$

$$\vec{e_t} \cdot \vec{ds} = 0 \text{ at the contour.}$$

The function $V(x, y)$ is characteristic of the type of a guided wave.

Complex Amplitude of a Guided Wave

Let the real value of the constant C for the normalized guided wave, characterized by the function $V(x, y)$, be designated by C_1. By definition, the term *complex amplitude of the guided wave* will be applied to the ratio:

$$D = \frac{C}{C_1} \cdot \tag{2}$$

The power transmitted by the guided wave is therefore:

$$E = D \cdot D^* \tag{3}$$

where D^* represents the conjugate value of D.

Transmission Tensor

Let us now suppose that there are n guided waves each characterized by functions $V_i(x, y)$ which are all different. When part of the power E_i circulating in the corresponding waveguide is distributed among the $n-1$ others, the arrangement of n guides is called a junction or branching circuit. Let A_i designate each component of the incident vector $\vec{A}$ with n components, representing the complex amplitudes of the n incident waves in an n-dimensional complex space. Similarly, B_i will be the component of the transmitted vector $\vec{B}$ representing the whole of the complex amplitudes of the transmitted waves. Because of the linearity with respect to the values of D of the solutions, we can write:

$$\vec{B} = \overline{\overline{T}} \cdot \vec{A} \tag{4}$$

where $\overline{\overline{T}}$ represents the tensor having for matrix the square table lay out

$$\begin{Vmatrix} T_{11} \cdots T_{i1} \cdots T_{n1} \\ T_{1j} \cdots T_{ij} \cdots T_{nj} \\ T_{1n} \cdots T_{in} \cdots T_{nn} \end{Vmatrix}. \tag{5}$$

Tensor $\overline{\overline{T}}$ will be called the transmission tensor.

Each of its terms T_{ij} represents the coefficient of transmission from guide i into guide j. The term T_{ii} represents the reflection coefficient R_i in the direction of the incident wave ($R_i \equiv T_{ii}$). The term coupling factor f_{ij} (or transmission equivalent between guides i and j) is applied to the ratio:

$$f_{ij} = \frac{E_{ij}}{E_i} = T_{ij} \cdot T_{ij}^*, \tag{6}$$

T_{ij}^* designating the conjugate quantity of T_{ij}. E_i is the power incident into guide i, and E_{ij} the power transmitted into guide j when all other components A_k ($k \neq i$) are zero.

Reciprocity Theorem

Let us suppose that the media contained within each guide are all isotropic (ϵ_i and μ_i are then scalar values differing in each medium). It can be shown[1] that the transmission tensor is then a symmetrical tensor.

$$T_{ij} = T_{ji}. \tag{7}$$

This property is independent of the choice of the planes of origin. Formula (7) involves two relations: the transmission equivalent (or coupling factor) is the same for the two opposite directions of energy transfer, $f_{ij} = f_{ji}$, and the transit time is also the same in both directions.

Eq. (7) constitutes the reciprocity theorem. It is important to note that this theorem is invalid when at least one anisotropic medium exists (ferrite or electronic plasma).

Study of the Transmission Tensor in Lossless Media

Let us suppose that the media under consideration (electronic plasmas, isotropic, or anisotropic material media) are without electromagnetic losses of any kind. Also, since the media may be anisotropic, the transmission tensor is not necessarily symmetrical and the relation $T_{ij} = T_{ji}$ between the tensor components will, in general, not be observed. However, because of the absence of electromagnetic losses, a certain number of relations must exist between the transmission vector components. This new form for the tensor will be sought by supposing, therefore, that all media carrying waves introduce no electromagnetic losses.

This case results when the symmetrical components of the electric and magnetic induction tensor are real and when their antisymmetrical components are purely imaginary. In isotropic media, ϵ and μ are real numbers.

[1] This demonstration is given in: C.R.A.S. Part 217, p. 677, 1943. Note by Messrs. Gutton and Ortusi, "On the form of the reciprocity theorem on hertzian waves." It is important to note that (7) applies to amplitudes D of (2) and not to the amplitude of the electric vector or even the amplitude C of (1) which can be different for the two senses of propagation if the waves of guides i and j are not similar. This point is an important subject of error in the application of the reciprocity theorem, in particular when the media are not the same even with similar waves.

Determination of the Necessary and Sufficient Relations Between the Transmission Tensor Components in Lossless Media

Let us designate by L_a the total incident power

$$L_a = \sum A_i \cdot A_i^* = \vec{A} \cdot \vec{A^*}$$

and by L_b the total power transmitted

$$L_b = \sum B_i \cdot B_i = \vec{B} \cdot \vec{B^*},$$

where $\vec{A^*}$ and $\vec{B^*}$ are the conjugate vectors of $\vec{A}$ and $\vec{B}$. L_a and L_b then represent the lengths of the vectors $\vec{A}$ and $\vec{B}$. The theorem of the conservation of energy, in the case of lossless media, shows that $L_a = L_b$. The transmission tensor is therefore an *orthogonal tensor* (preserving angles and lengths).

This condition is necessary and sufficient. A junction of n guides, without electromagnetic loss in the media, is therefore characterized by an orthogonal transmission tensor. The components of this tensor are thus connected by the classical n^2 relations which fix the changes of coordinates on rectangular axes.

Let $\overline{\overline{T}}^*$ and $\overline{\overline{T}}_t$ be the conjugate and transposed tensors of the tensor $\overline{\overline{T}}$ defined by their respective components T_{ij}^* and T_{ji} (tensor $\overline{\overline{T}}_t$ is therefore identical to tensor $\overline{\overline{T}}$ when the media are isotropic, but differs from it in the contrary case).

Writing that tensor $\overline{\overline{T}}$ is orthogonal, it will be seen that we have:

$$\overline{\overline{T}}_t \cdot \overline{\overline{T}}^* \equiv 1 \text{ or } \overline{\overline{T}}^* \equiv \overline{\overline{T}}_t^{-1}.$$

$\overline{\overline{T}}_t^{-1}$ is the inverse tensor of the transposed tensor of $\overline{\overline{T}}$. The components of $\overline{\overline{T}}_t^{-1}$ are given by Krammer's rule. We have:

$$T_{tji}^{-1} = \frac{M_{ij}}{\Delta}$$

where M_{ij} represents the minor of component T_{ij} of tensor $\overline{\overline{T}}$ and Δ the determinant of (5) formed by the components of $\overline{\overline{T}}$.

We thus obtain the n^2 equations which give the components of the conjugate tensor $\overline{\overline{T}}^*$.

$$T_{ij}^* = \frac{M_{ij}}{\Delta}$$

which can be written:

$$\Delta = \frac{M_{ij}}{T_{ij}^*} = \frac{(-1)^{i+j}\Delta_{ij}}{T_{ij}^*} \tag{8}$$

Δ_{ij} being the determinant obtained by suppressing column i and row j in (5).

It should be noted that (8) is valid for any type of wave and for any media, isotropic or not. Eq. (8) thus constitutes the necessary and sufficient conditions to produce lossless junctions. In particular, it comprises the following relations (obtained, for instance, by developing Δ along each row and column):

$$\sum_{j=1}^{j=n} T_{ij} \cdot T_{ij}^* = 1 \quad \text{whatever the value of } i.$$

$$\sum_{j=1}^{j=n} T_{ji} \cdot T_{ji}^* = 1 \quad \text{whatever the value of } i.$$

Δ has a modulus equal to 1. We can therefore write $\Delta = e^{in\phi}$; ϕ will be called the junction phase angle.

Eq. (8) is usually written in the equivalent form

$$(-1)^{i+j}\frac{\Delta_{ij}}{T_{ij}^*} = e^{in\phi} \quad \text{whatever the values of } i \text{ and } j \tag{9}$$

It will now be shown how (9) can very simply define the principal categories of junctions and the various functional properties of these junctions of guides.

Junctions of Two Guides

Isotropic media is given in (5), in this case, by making

$$T_{12} = T_{21} = T \begin{Vmatrix} R_1 & T \\ T & R_2 \end{Vmatrix}.$$

Eq. (9) becomes:

$$R_1 R_2 - T^2 = \frac{R_2}{R_1^*} = \frac{R_1}{R_2^*} = -\frac{T}{T^*} = e^{2i\phi}.$$

R_1 and R_2 having the same modulus, the planes of origin may be chosen to give the relations:

$$R_1 = R_2 = \cos u \cdot e^{i\phi}.$$

With respect to those planes, we have:

$$T = i \sin u \cdot e^{i\phi}.$$

u is the reflection angle $(-\pi/2 < u < \pi/2)$:

$0 < u < \pi/2$ corresponds to a capacitive obstacle;
$-\pi/2 < u < 0$ corresponds to an inductive obstacle.

It is easy to see, when the longitudinal dimension of the obstacle is small, that u is related to the normalized susceptance Z by:

$$Z = -\frac{i}{2} tgu$$

The measurement of u (and consequently of Z) is made, in the laboratory of the Compagnie Generale de TSF, according to the method of Fig. 1.

The obstacle is put at an approximately equal distance of $\lambda/2$ between a detector and a conducting wall. The frequency of the source is modified in order to have two zeros at the detector which gives the u angle by the two corresponding values of the electric distance θ, $\theta_1 = 0$ and $\theta_2 = u$. This is a very accurate method for measuring small values of the u angle of obstacles used for uhf filters of large surtension.

Nonisotropic Media-Gyrator

The tensor $\overline{\overline{T}}$ is no longer symmetrical when the media are not all isotropic. By definition, when $T_{12} = -T_{21}$ the two-way junction forms a *gyrator*. Eq. (5) becomes a gyrator with lossless media, by making

$$T_{12} = -T_{21} = T \left\| \begin{matrix} R_1 & T \\ -T & R_2 \end{matrix} \right\|.$$

Eq. (9) becomes:

$$R_1 R_2 + T^2 = \frac{R_2}{R_1^*} = \frac{R_1}{R_2^*} = \frac{T}{T^*} = e^{2i\phi}.$$

It is thus seen that, for planes of origin so chosen that $R_1 = R_2$

$$R = \cos u \cdot e^{i\phi}$$
$$T = \sin u \cdot e^{i\phi}.$$

A tuned gyrator ($R_1 = R_2 = 0$) corresponds to the case where $u = \pi/2$; we then have, making $\phi = 0$, the relations:

$$T_{12} = 1 \quad \text{and} \quad T_{21} = -1.$$

The incident wave, in one direction, suffers no change of phase and no reflection. In the other direction, it suffers a phase change of 180°, without reflection; this condition is equivalent to a transit time, modified by half a period.

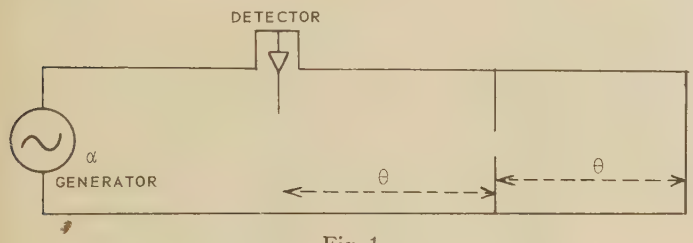


Fig. 1

Junction Among Four Guides

As an example of application of (9), the necessary and sufficient conditions will be sought which will give, in lossless media assumed to be isotropic, a matched directional coupling. The latter is characterized by the matching requirements: $R_1 = R_2 = R_3 = R_4 = 0$ and that of decoupling $T_{13} = 0$ (guides 1 and 3 being the decoupled guides). The matrix of tensor $\overline{\overline{T}}$ is:

$$\left\| \begin{matrix} 0 & T_{12} & 0 & T_{14} \\ T_{12} & 0 & T_{23} & T_{24} \\ 0 & T_{23} & 0 & T_{34} \\ T_{14} & T_{24} & T_{34} & 0 \end{matrix} \right\|.$$

Eq. (9) then becomes:

$$\Delta = \Delta^{1/2} \frac{T_{34}}{T_{12}^*} = \Delta^{1/2} \cdot \frac{T_{12}}{T_{34}^*} = \Delta^{1/2} \cdot -\frac{T_{23}}{T_{14}^*}$$
$$= \Delta^{1/2} \cdot -\frac{T_{14}}{T_{23}^*} = \frac{0}{T_{24}^*} = e^{4i\phi}. \qquad (10)$$

Designated by $f_{ij} = T_{ij} \cdot T_{ij}^*$ and by ϕ_{ij} the coupling factors and the arguments of the transmission coefficients T_{ij}, (10) becomes:

$$\begin{cases} f_{12} = f_{34} = \cos^2 w \\ f_{14} = f_{23} = \sin^2 w \\ f_{24} = 0 \\ \phi_{12} + \phi_{34} = {}_{\pi}2\phi \\ \phi_{14} + \phi_{23} = -\pi + 2\phi. \end{cases} \qquad (11)$$

Eq. (11), which is quite general, is valid whatever the form of the matched hybrid junction. In particular, it allows the following important theorem to be stated: matching of a hybrid junction, in which guides 1 and 3 are decoupled, *necessarily* involves complete decoupling between the other two guides 2 and 4.

The hybrid junction is necessarily formed by a series T and a shunt T of which the coupling factors are complementary. In the directional couplers, made with two distinct coupled guides, the phase angles of the directly transmitted wave and of the wave transmitted by coupling have a difference of $\pi/2$.

The Functional Properties of the Junctions with Anistropic Media

Anisotropic media provide certain functional properties in junctions which, from fundamental principles, are impossible of realization in isotropic media. In this section, two special cases will be examined.

Circulators

By definition a junction will be denominated a circulator when it consists of n guides defined by the following conditions regarding the components of its transmission tensor:

$$T_{ij} = 0 \text{ except for } j = i \text{ or } j = i + 1$$

with, of course, $T_{n1} \neq 0$ $(n+1 \equiv 1)$.

A circulator is shown in Fig. 2. For example, a 4-guide circulator will have a transmission tensor represented by the following:

$$\left\| \begin{matrix} R_1 & 0 & 0 & T_{41} \\ T_{12} & R_2 & 0 & 0 \\ 0 & T_{23} & R_3 & 0 \\ 0 & 0 & T_{34} & R_4 \end{matrix} \right\|.$$

It will be noted that the circulator must necessarily contain at least one anisotropic medium, since the transmission tensor is not symmetrical.

Matching Conditions for Circulators: It can be shown that, in the case of lossless media, a circulator must necessarily be matched.

Eq. (9) then becomes:

$$\prod R_i + \prod \theta_i = \frac{\prod_{j \neq i} R_j}{R_i^*} = \frac{\prod_{j \neq i} \theta_j}{\theta_i^*} = \frac{R_i \prod_{j \neq i+1 \text{ and } j \neq i} \theta_j}{0} \qquad (12)$$
$$= e^{in\phi}$$

making $T_{i,\ i+1}=\theta_i$ where the symbol II represents the product of the terms excluding terms whose numbers are specially indicated. These equations show that all R_i coefficients are zero. Some mistuning introduced in one guide of the junction necessarily introduces an undesired coupling with one of the guides numbered $j \neq i+1$. When the circulator is matched, all θ_i coefficients have a modulus equal to 1 and their phases satisfy the relation

$$\sum \phi_i = n\phi.$$

These conclusions do not apply in the case of a lossy medium.

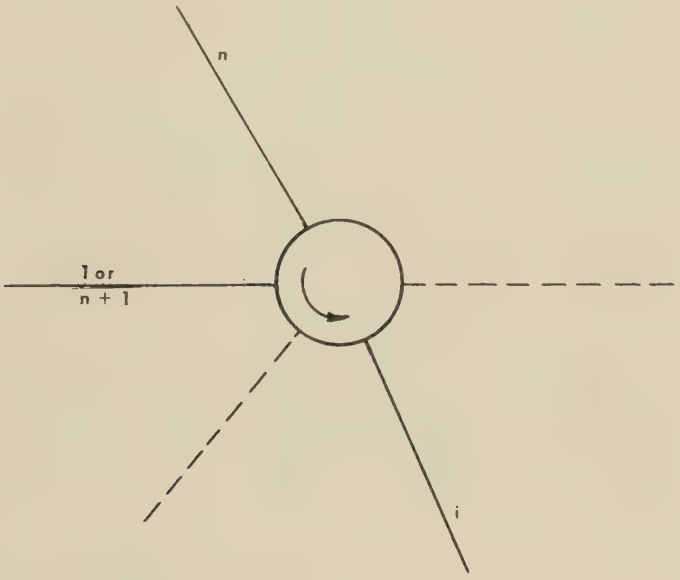


Fig. 2

Use of Circulators

Associated with band-pass filters, circulators are a very elegant means for effecting the junction of radio links. The diagram of Fig. 3 shows an arrangement with a single circulator with n guides.

Guide 1 contains the initial wave frequency modulated from the value $F=F_2$ up to the value $F=F_{n-1}$. Each guide i (with $2 \leqq i \leqq n-1$) terminates in an antenna or in a receiver tuned to frequency F_i. It contains a band-pass filter which passes, with a small reflection coefficient, the wave of frequency $F_i \pm \Delta F_i$ and reflects with a reflection coefficient close to 1 waves of frequencies F_j $(j \neq i)$.

Guide n terminates in an attenuation impedance and does not contain a filter. The mode of operation of the arrangement is immediately deduced from an examination of Fig. 3. No reflected wave can return to the input guide 1 whatever the quality of the filters F_i. Inadequate attenuation outside the passband produces linear cross talk in the transmitted wave (or the received wave, as the case may be). Too great a reflection within the passband produces an absorption of energy in the terminal impedance of guide n. No nonlinear cross talk need be expected if the modulation of the transmitter is satisfactory, since guide 1 is always matched. It is not essential to produce an n-section circulator in order to obtain the separation of n channels. A series of 4-section circulators can also be used.

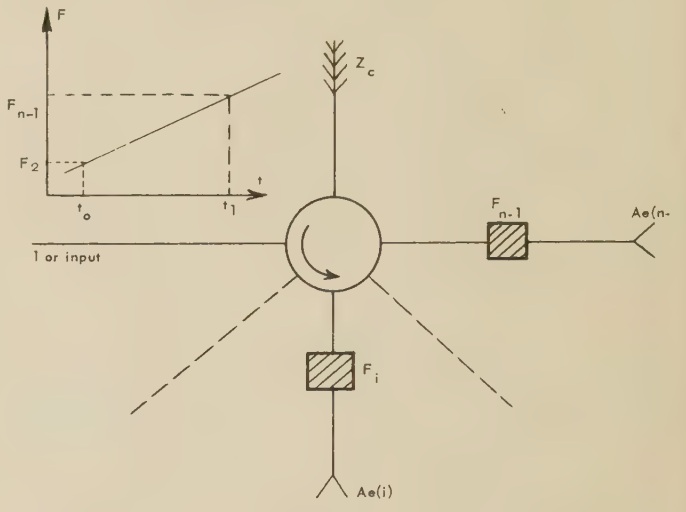


Fig. 3

Fig. 4 shows an example of an arrangement which is of special interest when the channels are grouped in pairs.

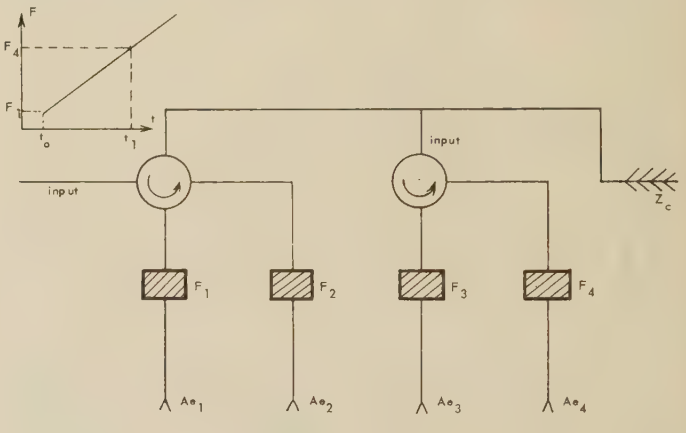


Fig. 4

Isolators

An isolator is a junction of two guides satisfying the following condition:

$$T_{21} = 0.$$

An isolator is represented by the diagram of Fig. 5.

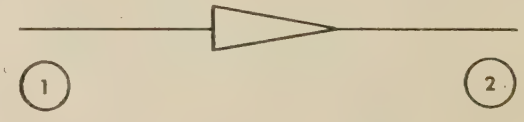


Fig. 5

It necessarily contains a nonisotropic medium except where T_{12} is also zero and where no coupling exists between guides 1 and 2 in either direction.

Necessary Presence of an Absorbing Body: It will now be shown that it must necessarily also contain an absorbing medium. If the medium is assumed to be lossless, (9) becomes in the general case of a junction of 2 guides of any sort:

$$\Delta = \frac{R_1}{R_2^*} = \frac{R_2}{R_1^*} = \frac{T_{12}}{T_{21}^*} = \frac{T_{21}}{T_{12}^*} = e^{2i\phi}.$$

From these equations, we deduce:

$$R_1 \cdot R_1^* = R_2 \cdot R_2^*$$
$$T_{12} \cdot T_{12}^* = T_{21} \cdot T_{21}^* \text{ or } f_{12} = f_{21}. \qquad (13)$$

Eqs. (13) provide the following theorem: In lossless media, even when anisotropic, the transmission equivalent and the modulus of the reflection coefficient are independent of the direction of energy transfer. No condition exists for transit times which may be different in the case of anisotropic media.

In particular, in the case of isolators, condition $T_{21}=0$ can only be verified if $T_{12}=0$, that is to say if no coupling exists between the two guides.

A lossless isolator cannot therefore be devised. This remark becomes generalized in the case of several guides. This theorem explains certain phenomena in guide junctions which appear to be paradoxical.

Consider, for instance, Fig. 6, which is a guide carrying wave H_{01} and a ferrite insert of thickness e and length l placed in the zone of the guide where the magnetic field is circularly polarized. It is well known[2] that, in this zone, the permeability tensor becomes a spherical tensor whose single component $\mu \pm \mu'$ depends on the direction of propagation. The applied continuous magnetic field and the ferrite insert may be so chosen that the effective magnetic permeability μ_e is equal to ϵ in the direction1 -2 and zero in the direction $2-1$.

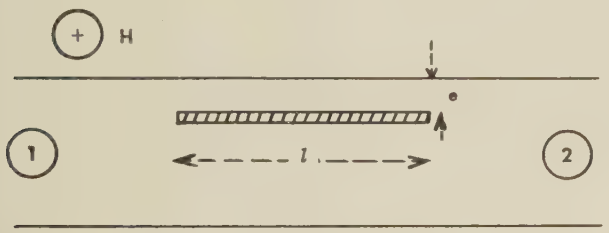


Fig. 6

In these conditions, if the length l is sufficient, we shall obtain the isolator conditions:

$$T_{21} = 0 \quad \text{with} \quad T_{12} \neq 1.$$

However, these conditions can only be attained with a suitable thickness e.

When thickness e tends to zero, no decoupling occurs between the alternating magnetic fields in the regions situated above the ferrite block and below it in the guide of Fig. 6. The coefficient T_{21} remains equal to 1. When the thickness e increases, these alternative magnetic fields are soon decoupled and, in the direction $2-1$ for which $\mu_e=0$, these fields then tend towards the value corresponding to the guide limited by the ferrite insert and the lower metal part.

When the frequency is sufficiently low, it can be seen that T_{21} then decreases very rapidly while T_{12} remains equal to 1. But the general conditions of the foregoing theorem impose that the decrease of T_{21} shall necessarily be attended by an absorption in the ferrite block proportional to the thickness e.

It is in fact this absorption which makes it possible to decouple the alternating magnetic fields in the two parts of the guide and, hence, to obtain a value of T_{21} less than that of T_{12} and which can even be zero.

With certain ferrite elements, it is possible to obtain the conditions of an isolator ($T_{12}=1$ and $T_{21}=0$) with the necessary losses remaining small. In that case, one can have $|R_1|=0$ and $|R_2|1$.

With that isolator and a conducting wall farther into the guide, one can realize a cavity whose surtension is nearly independent of the frequency passband. That device constitutes an energy accumulator which can have several applications. For instance, in Fig. 7, in-

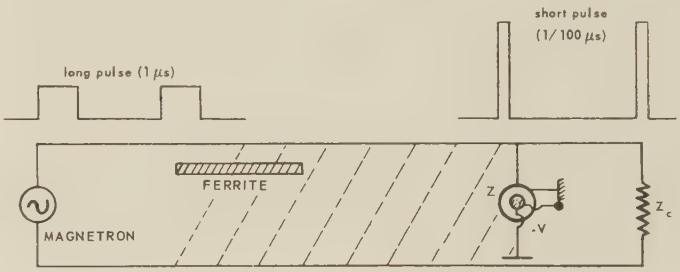


Fig. 7

stead of the conducting wall, it is possible to put an impedance Z formed by an ionic or electronic commutator; *i.e.*, a device whose impedance is modified by the voltage V. For instance, we have

$$V \ll V_0 \rightarrow Z = 0$$
$$V > V_0 \rightarrow Z = \infty.$$

When V is less than V_0, the cavity formed between the isolator and the impedance Z can be excited by a magnetron acting with relatively long pulses (1 μs for instance).[3]

When the voltage V is greater than V_0, the cavity is open and one obtains, in the characteristic impedance Z_c of the guide, pulses of large amplitude and very short time (10^{-3} μs) when the isolator losses remain small.

Use of Isolators: Isolators can be used in two different ways:

[2] J. R. Roven, "Ferrite in microwaves applications," *Bell Syst. Tech. Jour.*, vol. 32, p. 1333; November, 1953.

[3] R. R. Warnecke and J. Ortusi, U. S. Patent No. 2,543,213.

Associated to an oscillator, isolators prevent frequency drift caused by waves reflected from the antenna. This advantage is specially valuable for a frequency modulated klystron.

Similarly, in the case of its use on measuring benches where the isolation factor of the isolator (ratio T_{12}/T_{21}) expressed in decibels, does not need to be particularly high. From the point of view of reaction on the transmitter an isolator of 20 decibels is equivalent to an isotropic attenuator of 10 decibels but has the advantage of almost fully maintaining the power of the generator.

Introduced in a delay line, isolators avoid, in traveling wave tubes, coupling between the input and output of the tube. Such coupling may arise either through a mismatch in the output guide, or through the presence of a wave whose direction of propagation is opposite to that of the circulation of energy. The use of an attenuating impedance with isotropic medium, uniformly distributed or not, can suppress spurious oscillations only at the expense of the tube efficiency and of a deterioration of the gain characteristic with frequency. The use of an isolator, comprising an anisotropic body in the delay line, avoids these disadvantages.

Phenomenological Vector Model of Microwave Reflection from the Ocean*

C. I. BEARD†, I. KATZ†, AND L. M. SPETNER†

Summary—**A model of one-way transmission of microwave electromagnetic signals over the ocean surface is developed from experiment. The received signal is described as a vector sum of a constant direct signal, a coherent reflected signal, whose amplitude and phase are fixed by geometry and sea state, and a fluctuating reflected component of random amplitude and phase. By interpreting experimental data in the light of this phenomenological model it has been possible to relate, quantitatively, the coherent and incoherent reflected signal and total signal to geometry and sea state. The results give support to the theoretical expression previously derived by Ament and others relating the coherent reflected signal to "apparent ocean roughness." In addition, the general shape of the curve relating the incoherent scattering to "apparent ocean roughness" has been established and its asymptotic value found.**

Introduction

IN ONE-WAY transmission of microwave electromagnetic signals over the ocean between transmitting and receiving points within the line of sight, the received electric-field strength can be represented as the vector sum of a direct signal and a water-reflected signal. The resultant is a random, time-varying signal which is a function of geometry and ocean-surface parameters, or "sea state." Theoretical treatments of this forward scattering have been limited to special cases[1] or are formal solutions which have not been reduced to specific functional relations.[2–4] Two recent reports more closely related to the present work and which derive explicit relations for the coherent reflected term are those by Ament[5] and Goldstein and Goldmuntz.[6] In addition, extensive unpublished theoretical work has been done at the Applied Physics Laboratory by Spetner.

On the other hand, past experimental work has been regrettably deficient in that the ocean surface fluctuations have not in general been measured. It is only in the past few years that sea surface measurements have been included in propagation experiments to provide the data necessary for formulating an experimental description of forward scattering. The several experiments providing the data that made possible the analysis presented in this paper were performed by the Electrical Engineering Research Laboratory (EERL) of the University of Texas. EERL and the Applied Physics Laboratory of the Johns Hopkins University have been associated since the beginning of 1951 in studies of microwave reflection from the water.

It is the purpose of this paper to report a phenomenological vector model of reflection which has been induced from these experimental data and which quantitatively relates the coherent and incoherent reflected signals and the total signal to geometry and sea state. This representation ties together many diverse experiments into a consistent pattern which has heretofore been lacking.

This paper will describe the phenomenological model by first giving its vector representation and postulates, followed by the experimental supporting evidence for its validity. A homogeneous atmosphere will be assumed throughout.

* Manuscript received by the PGAP, August 11, 1955; revised copy received, February 8, 1956.

† Applied Physics Lab., Silver Spring, Md.

[1] S. O. Rice, "Reflection of electromagnetic waves from slightly rough surfaces," *Comm. on Pure and Applied Math.*, vol. 4, pp. 351–378; August, 1951.

[2] Lord Rayleigh, "On the dynamical theory of gratings," *Proc. Roy. Soc. A*, vol. 79, p. 399; 1907.

[3] C. T. Tai, "Reflection and refraction of a plane electromagnetic wave at a periodical surface," Harvard Tech. Report No. 28, Cruft Laboratory; January 15, 1948.

[4] V. Twersky, "Multiple scattering of radiation by an arbitrary configuration of parallel cylinders," *J. Acoust. Soc. Am.*, vol. 24, pp. 42–46; January, 1952. (More complete bibliographies are given in the papers of Rice and Twersky.)

[5] W. S. Ament, "Toward a theory of reflection by a rough surface," PROC. IRE, vol. 41, pp. 142–146; January, 1953.

[6] H. Goldstein and L. A. Goldmuntz, "Mean square signal computation," Nuclear Development Associates Report NDA-18-2; August 18, 1952.

Vector Model and Postulates

If electromagnetic radiation is transmitted over a calm ocean and the resulting field is probed vertically, a well-defined interference pattern will be found, indicating a strong coherent reflection from the surface. If this experiment is repeated as the ocean becomes rougher, the interference pattern will be less pronounced. With very rough surfaces there will be no identifiable interference pattern, indicating incoherent scattering is predominant. Attempts to describe this process analytically led to the following assumption:

Postulate 1

The total signal (T) is the vector sum of (Fig. 1):

1) the direct ray (D) of constant magnitude and phase,
2) a coherent reflected ray (C) whose amplitude and phase are fixed by a given geometry and sea state, and
3) a fluctuating or incoherent reflected component (I) of random amplitude and phase.

(The quantities T, D, C, and I represent electric field strengths.)

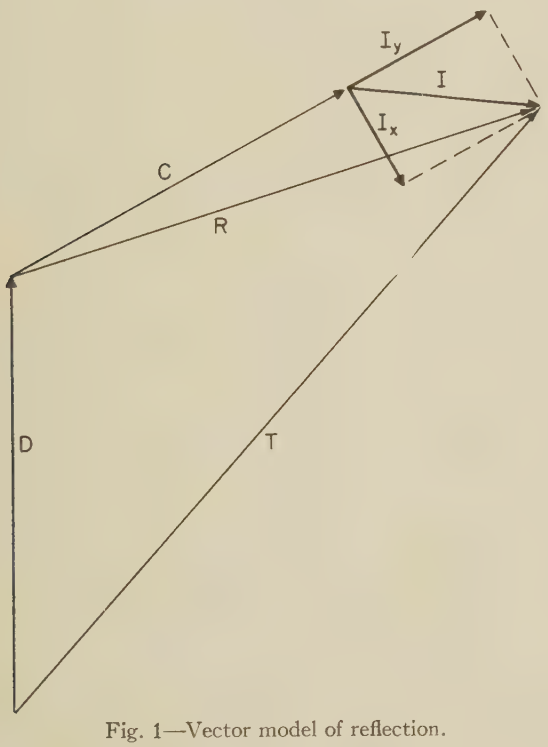


Fig. 1—Vector model of reflection.

If the vector I is assumed to be the resultant of individual vectors from a large number of independent random scatterers, following Goldstein's treatment of backscattering from the sea,[7] then it follows:

Postulate 2

The incoherent vector I is the resultant of two orthogonal independent vectors each representing a Gaussian random process with zero mean. (For convenience, the directions of the two orthogonal vectors are taken along and normal to C and their standard deviations and spectra are assumed to be equal.)

Description of Experiments

This paper draws on experiments performed by associated laboratories in the course of investigations into propagation problems. These experiments were quite varied in nature using several frequencies in the microwave region; all involved only one-way propagation. Certain facts were gleaned from each of the experiments which, viewed together in the light of the model, form a consistent pattern. Since these investigations form the basis for the data a necessarily short résumé of each will now be presented. More details may be found in original reports.

Lake Austin

In February, 1952, the Electrical Engineering Research Laboratory (EERL) of the University of Texas performed an experimental study of the interference patterns existing over a relatively smooth body of water, Lake Austin, Texas.[8] The transmitter was placed near the water's edge on one side of the lake while continuous measurements of field strength were made by a receiver mounted on a 50-foot elevator across the lake, a distance of 1,430 feet. (Such experiments are generally referred to as "height-gain measurements.") Three wavelengths were used: 9, 3.2, and 0.86 centimeters.

Gulf of Mexico

In August, 1952, EERL carried out a series of propagation studies over two paths in the Gulf of Mexico to measure the reflection characteristics of ocean water at radio wavelengths of 9.0, 5.3, 3.2, and 0.86 centimeters.[9–11] One path was across Barataria Pass at Grand Isle, La., a distance of about 2,800 feet, and the other was between two oil drilling platforms 5,000 feet apart, eight miles offshore. Height-gain measurements were made at Barataria Pass whereas fixed-point measurements over a time interval of a few minutes near interference maxima were made on the platform path. Some measurements were also made at 3.2 cm with dual antennas at the receiver end.

San Clemente Island

A continuous interference pattern was measured by EERL in October, 1953 using an airplane-mounted transmitter and a stationary 3-cm vertically-polarized

[7] D. E. Kerr, "Propagation of Short Radio Waves," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 553–554; 1951.

[8] A. W. Straiton, J. R. Gerhardt, A. H. LaGrone, and C. W. Tolbert, "Reflection of Centimeter and Millimeter Radio Waves from the Surface of a Small Lake," EERL, Univ. of Texas, Report No. 63; May 15, 1952.

[9] A. H. LaGrone and C. W. Tolbert, "Reflection Studies of Millimeter and Centimeter Radio Waves for Gulf of Mexico Paths," EERL, Univ. of Texas, Report CM-753; October 31, 1952.

[10] W. J. McKune and H. W. Smith, "Comparison of Power Spectrum Estimates of Overwater Microwave Radio Signal and Associated Water Waves," EERL, Univ. of Texas, Report No. 68; May 6, 1953.

[11] A. H. LaGrone, A. W. Straiton, and H. W. Smith, "Synthesis of Radio Signals on Overwater Paths," EERL, Univ. of Texas, Report No. 71; April 30, 1954.

receiver 45 feet above the shoreline on San Clemente Island, California. The airplane flew toward the receiver at a constant altitude of 125 feet starting at a range of 10,000 yards.[12]

Golden Gate

Further over-water propagation measurements were made by EERL in the spring of 1953 over the Golden Gate at San Francisco, California.[11] The geometries of the two over-water paths used are summarized in Table I. Heights are above mean lower low water.

TABLE I

Path	Receiver Height (feet)	Transmitter Height (feet)	Range (feet)
I	186	85	16,000
II	186	260	16,500

The paths extended across a water surface approximating open-ocean conditions. This experiment was quite varied in nature; it included measurements of total signal and signal with a null placed on the specular reflection to obtain the direct signal magnitude for normalization; simultaneous reception was accomplished on two antennas separated vertically by 50 and 100 wavelengths alternately. Extensive ocean-wave and meteorological measurements were made.[13]

San Clemente, Convair

An experiment similar to the one previously described under the EERL San Clemente Island Tests was carried out by Consolidated Vultee Aircraft Corporation at San Clemente Island, California, during May, 1954.[14] The airplane runs were made at several constant altitudes on incoming courses starting at ranges of about 7,500 yards. A 3-cm transmitter using a wide antenna pattern was carried in the airplane; reception was accomplished at a 200-foot-altitude stationary point at the shoreline.

Experimental Evidence Supporting Phenomenological Model

From the experiments just described data were obtained which are now offered in evidence of the model's validity. The initial clue was in the form of numerical values of the Rayleigh criterion of roughness. When it was clear that a transition zone existed between smooth and rough surfaces as measured by the parameter $h\psi/\lambda$, a natural question was raised whether this parameter was indeed a quantitative measure of surface roughness; the present section will show data supporting this. (The quantity h is here defined as the rms water-surface fluctuation measured from the mean water level; ψ is the grazing angle; and λ is the electromagnetic wavelength.) From the form of the height-gain runs came a suggestion that incoherent noise increased with the same parameter; the experimental data led to the quantitative relation between I and $h\psi/\lambda$.

Transition Region

Until recently the best estimate of the Rayleigh smooth-to-rough transition in the microwave region was given by Kerr as $h\psi/\lambda = 25$,[15] where ψ is in milliradians. Evidence from height-gain and airplane runs at Lake Austin, Barataria Pass and San Clemente Island indicated transition values of the roughness parameter $h\psi/\lambda$ between 8 and 60. Although at first it would seem that this is a large spread it should be noted that the values bracket the critical value given by Kerr. Furthermore, the random nature of the incoherent component makes the selection of a transition point quite difficult. Rather, the change from a smooth to a rough surface should be thought of as a zone of transition, not a point.

Of the three experiments giving evidence on the transition region, the Gulf of Mexico set was of most interest. In this set there were three different frequencies involved with both horizontal and vertical polarization. Yet a roughness parameter value of about 10 ± 2.5 was selected for all runs. The variation was indeed small considering the qualitative nature of what is called a transition.

Curve of Coherent Term

The term "reflection coefficient" is generally taken to mean the ratio of coherent reflected ray to that of the direct ray, C/D. Since C/D decreases with increasing ψ even for a smooth sea, in the following discussion this ratio will be divided by the smooth-sea reflection coefficient r. One point on the curve of the coherent term vs effective roughness has already been established,[16,8] namely, $C/Dr = 1$ for $h\psi/\lambda = 0$. Values of the coherent term for $h\psi/\lambda > 0$ were obtained from the San Clemente (EERL) airplane run and the Golden Gate experiments.

The continuous interference pattern obtained at San Clemente Island (EERL) was over a sufficient range of roughness to yield several values for the coherent term. They were obtained by drawing smooth curves through the maxima and minima of the pattern in the region of interest. Apparent reflection coefficients were computed from the upper and lower smoothed curves. These are plotted in Fig. 2. No dependable points were obtained for higher values of $h\psi/\lambda$ because the smoothing process

[12] Since the pattern referred to above was obtained at APL's request only as an incidental sideline to other activity and was worked up at APL, no published reference is available.

[13] K. H. Jehn, J. R. Gerhardt, D. F. Metcalf, and S. J. Prosser, "Some Meteorological and Oceanographic Characteristics of the Golden Gate, California, Area," EERL, Univ. of Texas, Report CM-760; February 26, 1954.

[14] Records from these runs were made available to APL for reduction and analysis; no published reference is as yet available for this experiment.

[15] This number was obtained by taking Kerr's obstacle height as $H_{1/10} = 5.1\ h$. For this transformation see M. S. Longuet-Higgins, "On the statistical distribution of the height of sea waves," *Journal of Marine Research*, vol. 11, p. 245; December, 1952.

[16] D. E. Kerr, *op. cit.*, pp. 402, 423.

is no longer valid when the incoherent term becomes too large. The ordinates of the two points are estimated to have an accuracy of ±20 per cent. Because of the uncertainty in wave height, the abscissa position is known to only ±50 per cent. The solid curve is one previously derived from theoretical considerations by others [5,6] and has the analytic form

$$\frac{C}{Dr} = \exp - 8(\pi h\psi/\lambda)^2.$$

In this equation ψ is in radians; everywhere else in this paper ψ is in milliradians.

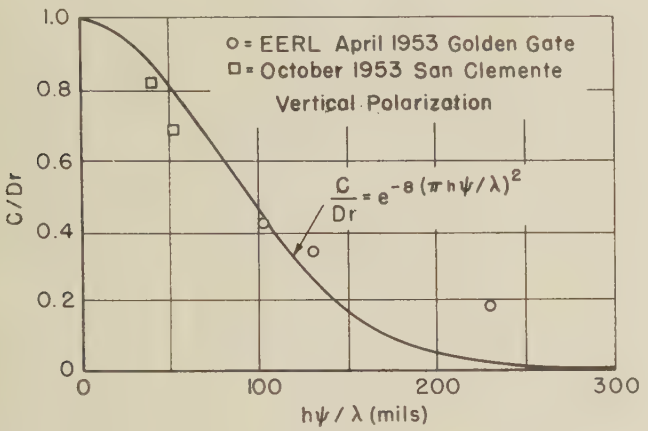


Fig. 2—Coherent term, theoretical and experimental.

Additional points for higher $h\psi/\lambda$ on this graph were obtained from the Golden Gate experiment. These are the three points plotted between values of 100 and 250 in $h\psi/\lambda$. These were obtained from the records of total signal at three antennas, recorded two at a time.

The following analysis describes the method used for extracting the coherent term from simultaneous records of total signal without knowing one's position in the interference pattern. According to the phenomenological model (Postulate 1) the total signal may be written

$$T^2 = (D + C\cos\phi + I_y\cos\phi - I_x\sin\phi)^2 + (C\sin\phi + I_y\sin\phi + I_x\cos\phi)^2,$$

where ϕ is the angle the coherent vector makes with the direct ray vector and I_y and I_x are components of the incoherent vector along and normal to C. Assuming the mean values of I_y and I_x are zero (Postulate 2) and that ϕ does not change appreciably from atmospheric fluctuations or water level changes during the 10-minute recording, $\overline{T^2} = D^2 + C^2 + 2DC\cos\phi + \overline{I^2}$. The phases at the three antennas, spaced vertically may be given as

$$\phi_c = \frac{4\pi h_t}{\lambda R}\, h_c = qh_c$$

$$\phi_b = qh_b \text{ and } \phi_a = qh_a$$

where h_t is transmitter height and h_a, h_b, and h_c are receiver antenna heights, respectively. The mean total signal at each antenna is given by:

$$\overline{T_a^2} = D^2 + C^2 + 2DC\cos(\phi_c + 2\Delta) + \overline{I_a^2}$$

$$\overline{T_b^2} = D^2 + C^2 + 2DC\cos(\phi_c + \Delta) + \overline{I_b^2}$$

$$\overline{T_c^2} = D^2 + C^2 + 2DC\cos\phi_c + \overline{I_c^2}$$

where

$$\Delta = q(h_b - h_c) = q(h_a - h_b).$$

The assumption is made that $\overline{I^2}$ is identical at all three antennas, a reasonable assumption for the geometry involved in this experiment. Now

$$\overline{T_a^2} - \overline{T_c^2} = 2DC[\cos(\phi_c + 2\Delta) - \cos\phi_c] \quad (1)$$

$$\overline{T_b^2} - \overline{T_c^2} = 2DC[\cos(\phi_c + \Delta) - \cos\phi_c). \quad (2)$$

The value of the direct ray was measured in the Golden Gate experiment by using a binomial-array transmitter to place a broad null in the specular reflection direction. The value of Δ is computed from geometry. Thus there remain two unknowns C and ϕ_c which are now determinable from the two equations. This method of obtaining the coherent term is quite general and is independent of the statistics of the incoherent term. Its reliability is dependent on how accurately D, $\overline{T^2}$ and the geometry are determined. $\overline{T_a^2}$ and $\overline{T_b^2}$ were independently assumed to be different from their measured values by up to ±20 per cent (corresponding to the 10 per cent rms error found in D) and the resulting errors in C were calculated. This represents the most pessimistic range of error since some correlation was observed between D_a and D_b. The greatest error in Δ is caused by the possible error of a transit sighting on the transmitter position. From surveying texts this was estimated as ±1 minute and the rms assumed to be ½ minute. The error in C/Dr from this source is independent and was added to the first by the square root of the sum of the squares to form a resultant worst possible rms error of ±100 per cent in the values of C/Dr.

Incoherent Component

The fact that, as $h\psi/\lambda$ increases, the incoherent noise term I increases was based on examination of the data from the various experiments already mentioned. The point at $h\psi/\lambda = 0$ is immediately established by the absence of "noise" on height-gain runs over a smooth lake surface.[8] Before presenting further experimental data for $h\psi/\lambda > 0$, a derivation is necessary to indicate the method of obtaining I from total signal-strength measurements.

First, for simplicity, let $\phi = 0$ so that C is in phase with D, and let $D_1 = D + C$. Then from Postulate 1:

$$T = [(D_1 + I_y)^2 + I_x^2]^{1/2}.$$

Assuming

$$\frac{I_x^2}{(D_1 + I_y)^2} \ll 1,$$

a series expansion can be made for T. Expanding the individual terms in series and neglecting terms of higher order than I_x^4, one obtains:

$$T \cong (D_1 + I_y) + \frac{I_x^2}{2D_1}\left(1 - \frac{I_y}{D_1} + \frac{I_y^2}{D_1^2} - \cdots\right) - \frac{I_x^4}{8D_1^3} + \cdots .$$

Making the following three assumptions of Postulate 2, that (a) $\overline{I_y} = \overline{I_x} = 0$, (b) $\overline{I_x^2} = \overline{I_y^2}$, and (c) I_y and I_x are Gaussian (then $\overline{I_x^4} = 3\sigma^4$, where $\overline{I_x^2} = \sigma^2$), one obtains:

$$\overline{T}^2 \cong D_1^2 + \sigma^2 + \frac{\sigma^4}{2D_1^2} + \cdots .$$

Now $\overline{T^2} = D_1^2 + 2\sigma^2$ under assumptions (a) and (b) above. Let T_{sd} be the standard deviation of T. Then,

$$T_{sd}^2 = \overline{T^2} - \overline{T}^2 \cong \sigma^2 - \frac{\sigma^4}{2D_1^2} .$$

Since, as will be shown later, σ/Dr does not exceed 0.33, and since, at a maximum in the interference pattern D_1 cannot drop below D, the largest error in T_{sd}^2 arising from the approximation $T_{sd} \cong \sigma$ is 3 per cent.

In any experiment, therefore, the standard deviation of the total signal at or near an interference maximum is taken as the standard deviation of one of the two orthogonal components (I_y or I_x) of the incoherent noise vector I.

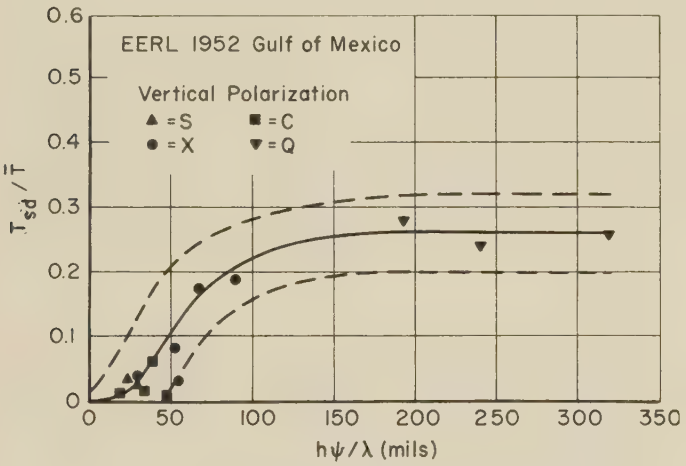


Fig. 3—Experimental values of incoherent term showing trend with $h\psi/\lambda$.

Let us first examine data from the earlier Gulf experiments (Fig. 3) that led up to the final results, primarily to show the trend with ocean roughness. Fig. 3 gives the complete plot of the vertical-polarization data for all values of $h\psi$ and λ but plotted as a function of $h\psi/\lambda$ for the 1952 Gulf of Mexico experiments.[17] An estimate of the uncertainty bounds of the data, taking into account the wave uncertainty and the spread in the radio data, is given by the dashed lines. The solid line is a smooth curve drawn through the points. The trend shown in this curve is clear. For low values of $h\psi/\lambda$ the fluctuating component is near zero, and at high values of $h\psi/\lambda$ the curve levels off. Since the direct ray was not measured, however, no normalization to D is possible for these points. It may be noted, however, that normalizing to D and dividing by r would result in a steeper slope at low values of $h\psi/\lambda$. Analysis of horizontal-polarization data obtained in the Gulf of Mexico[9] showed the same trend with h/λ as shown in Fig. 3.

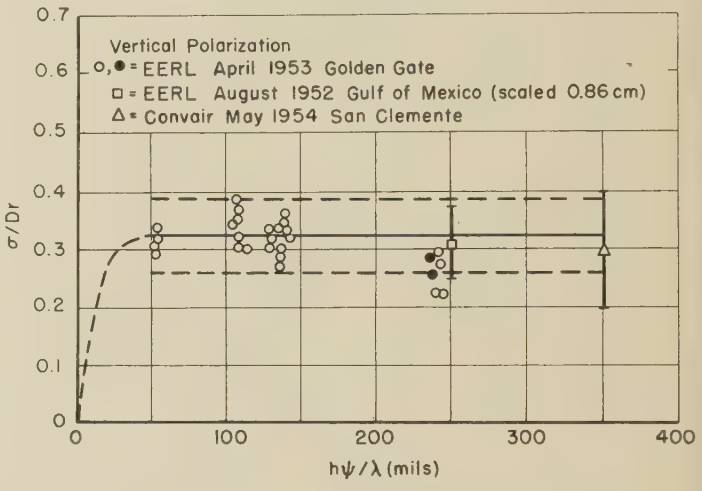


Fig. 4—Normalized values of incoherent term from experiment.

The latest and most reliable results concerning incoherent noise and for which normalization was possible are given in Fig. 4. At the lowest values of $h\psi/\lambda$ reached in these data the value of σ has already leveled off. One may only conclude that the transition region for "smooth" to "rough" for the incoherent energy lies below $h\psi/\lambda = 50$. For the points denoted by open circles the direct-ray strength was measured by using a binomial-array transmitter which placed a null in the specular-reflection direction. For the other points, D was not measured separately, but at these high values of $h\psi/\lambda$, $C \to 0$, and $\overline{T} \cong D$, so that these points are approximately normalized. The circled points represent data samples of ten minutes in length and the other points are from samples of variable lengths less than this.

Another source of incoherent noise information is the measurement of the signal received from an airplane transmitter at large values of $h\psi$. The normalization approximation is used that $\overline{T} \cong D$. In this case, however, $\overline{T}$ is varying as $1/R$ and hence in computing the standard deviation each departure from the mean has to be divided by the mean at that point. In order to get enough points to compute T_{sd}, it is preferable to use a very short section of each airplane run (to limit the variation in grazing angle) and to use many such runs to obtain a reliable value of T_{sd}. Four airplane flights

[17] Only four of the points on Fig. 3 are given in reference 9; the other points have been worked up from EERL's original data. On August 7, 1952, the wave-gage accuracy was open to question and a best estimate from all data was made that $h \cong 0.5 \pm 0.25$-feet rms.

made by Convair off San Clemente Island on May 4, 1954 were analyzed in this manner. The assumption in using these runs from a moving point is that the magnitude of the incoherent noise is the same as in the static case and that only the spectrum may change. A summary of the data is given in Table II below and the average is plotted on Fig. 4. The solid line is drawn at the level of the average of all the points. It is extended arbitrarily to zero by a broken line. The horizontal dashed lines indicate $\pm$ twice the rms error or the ± 20 per cent bounds about the solid line.

TABLE II

Run No.	Airplane Height (feet)	$\frac{T_{sd}}{T}$	$h\psi/\lambda$ (mils)	r	$\frac{T_{sd}}{Dr}$
7	50	0.15	300 to 400	0.45 to 0.33	—
11	50	0.13	300 to 400	0.45 to 0.33	—
6	100	0.07	300 to 400	0.45 to 0.33	—
9	150	0.10	300 to 400	0.45 to 0.33	—
Ave.	—	0.11 ± 0.04	300 to 400	0.39	0.3 ± 0.1

Conclusion

By using a vector representation of the total, direct, coherent-reflected and incoherent-reflected signals it has been possible to arrive at quantitative relationships between these signals and sea state:

a) Experimental evidence is given in support of a theoretical expression previously derived by other workers of the relation of the coherent reflected signal to "apparent ocean roughness."

b) The general shape of the curve of the incoherent-reflected term to "apparent ocean roughness" has been established and quantitative values have been assigned.

Although this vector representation may be useful in predicting certain phenomena related to forward scattering over smooth- and rough-water surfaces, it is realized that it is only a beginning. All the experiments mentioned in this paper were performed before the vector model was formulated and are only partially suitable. New experiments are needed to explore concepts suggested by the model. (Some experiments pointed in this direction were accordingly planned and were performed in April and May of 1955 by EERL under the sponsorship of APL.) For example, it is highly desirable to obtain information regarding the statistics of I. One method for obtaining such data is to measure rf phase as well as amplitude. In addition, in the plots of C and I vs $h\psi/\lambda$ insufficient data are available to establish any dependence upon water-wave direction. There is reason to believe that a wave-direction effect would exist since in the special case of short paths with the propagation direction perpendicular to the wave direction as at Barataria Pass,[9] long swells may lift the entire Fresnel zones up and down. This effect should decrease as the radio-propagation direction approaches parallelism with the wave-propagation direction. Also for longer paths and short-crested waves the Fresnel zones are more likely to include a number of waves rather than just one, even when the radio- and water-wave-propagation directions are perpendicular.

Likewise, all types of "seas" have been lumped together in the parameter h, the rms water-surface fluctuation. For short ranges and long-period swells there may be only a few swells in the illuminated region. Since the statistics assumed for the incoherent term in Postulate 2 depend upon the existence of at least 5 or 10 scatterers in the reflecting area, one would expect a departure from the simple model under these conditions.

Refinements such as the effect of wave direction and the number of scatterers are suggested for future investigations. What has been attempted in this paper is the development of the simplest general model that would tie together presently available experimental facts.

Report on Comparative 100 MC Measurements for Three Transmitting Antenna Heights*

A. P. BARSIS† AND R. E. McGAVIN†

Summary—**This report evaluates measurements taken during August, 1952, at a frequency of 100 mc as transmitted from three transmitting sites at elevations ranging from 6,220 feet to 14,110 feet above mean sea level. Six receiving sites were used ranging in distance from approximately 50 miles to 620 miles from the transmitter sites. Results are presented in terms of hourly medians of recorded field intensity and their distributions, as well as the over-all median values and deviations derived from these distributions.**

Introduction

DURING JULY and August of 1952 mobile recording equipment was set up near Anthony, Kan., and Fayetteville, Ark., in order to investigate radio wave propagation over the Cheyenne Mountain path during a typical summer month. This procedure served to supplement the regular Cheyenne Mountain Field Station recording program[1] which provides for continuous recordings at four receiving sites at frequencies ranging from 92 to 1,046 mc. The acquisition of a mobile 1 kw transmitter operating at a frequency of 100 mc provided a possibility of comparing propagation from two 100 mc transmitting sources located at different ground elevations, and operating during alternate hourly periods. Accordingly, such alternate transmissions were made during the period August 6 to August 13, using the mobile transmitter at Camp Carson which is lower than the regular fixed 100 mc installation, and during August 14 to August 20, using the mobile transmitter at a location on top of Pikes Peak which is much higher than the fixed installation. It is the purpose of this paper to evaluate the results of these measurements.

Transmitting and Receiving Stations and Transmission Paths

Tables I and II opposite give pertinent data on the antennas and paths used for these measurements. Special rhombic antennas were erected for transmitting at Camp Carson, and for receiving at Anthony and Fayetteville, and a five-element horizontal Yagi array was built for the Pikes Peak site. The Pikes Peak installation including the transmitter truck is shown in Fig. 1. The 90-degree corner reflector used on Cheyenne Mountain is fed by a folded dipole located 0.28 wavelength from the corner of the reflector.

Pattern and gain measurements on the Camp Carson antenna were made by comparing the field produced by it with the field produced by a dipole mounted on the

Fig. 1

rear pole of the rhombic, the measurements being made with mobile receiving equipment 15 to 20 miles distant.

A similar procedure was used to determine the gains and patterns of the Anthony and Fayetteville receiving rhombics, except that the mobile unit was used for transmitting, and the signals were received on the rhombic, and on the auxiliary dipole mounted on the rear pole of the rhombic (see Fig. 2). In all three cases,

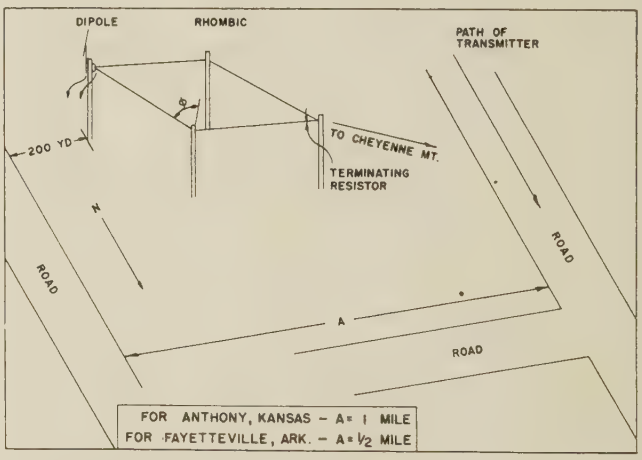


Fig. 2

tests were also made on the dipole antenna with the rhombic antenna lowered, to ascertain the effect of the rhombic on the gain of the dipole, so as to obtain the corrections therefor. For the two receiving rhombics, these tests were made only along the main axis of the rhombic, and the correction so derived was assumed constant over the range of azimuth angles used in the experiments. Fig. 3 shows a diagram of the antenna matching section used for the receiving rhombics.

* Manuscript received by the PGAP, February 8, 1956.
† National Bureau of Standards, Boulder, Colo.
[1] "Tropospheric propagation research," *NBS Tech. News Bull.*, vol. 36, pp. 120–124; August, 1952.

TABLE I

PERTINENT DATA—TRANSMITTING AND RECEIVING SITES

a) Transmitting Sites

	Location		Elevation of Antenna Above Mean Sea Level	Type of Antenna
	N. Lat.	W. Long.		
Cheyenne Mt. Summit	38.764°	104.864°	8,805 feet	Corner Reflector
Camp Carson	38.696°	104.830°	6,258 feet	Rhombic
Pikes Peak	38.839°	105.042°	14,115 feet	Yagi

b) Receiving Sites

	Location		Elevation of Antenna Above Mean Sea Level	Type of Antenna
	N. Lat.	W. Long.		
Kendrick	38.569°	103.984°	5,279 feet	Dipole
Karval	38.632°	103.572°	5,079 feet	Dipole
Haswell	38.383°	103.141°	4,334 feet	Dipole
Garden City	37.833°	100.858°	2,874 feet	Dipole
Anthony	37.240°	97.898°	1,375 feet	Rhombic
Fayetteville	36.107°	94.107°	1,385 feet	Rhombic

c) Receiving Sites—Distances from Transmitters (Miles)

	Cheyenne Mt.	Camp Carson	Pikes Peak
Kendrick	49.3	46.6	60.0
Karval	70.2	68.0	80.5
Haswell	96.6	93.8	107.4
Garden City	226.5	223.6	237.1
Anthony	393.5	390.7	404.1
Fayetteville	617.7	614.0	628.1

In the case of the Yagi antenna, the pattern measurement was made at Camp Carson, by rotating the antenna and measuring the relative signal strengths received from Cheyenne Mountain. The gain was measured at the maximum of the main lobe by using it as a transmitting antenna and comparing its field at 25 miles with that of a dipole set up in place of the Yagi.

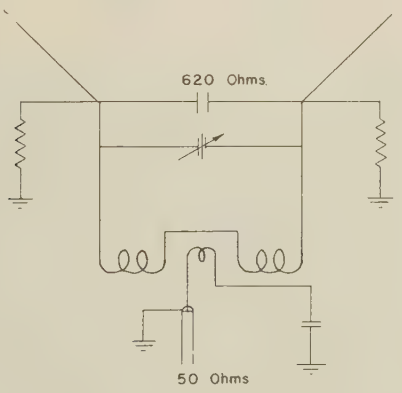


Fig. 3—Impedance matching section, 100 mc rhombic antennas, Anthony and Fayetteville.

All antenna gain data are given in decibels relative to an isotropic radiator under the assumption that all dipoles used either for comparative gain measurements, or as receiving antennas, compared closely to the ideal case as far as gain and pattern data were concerned.

TABLE II

ANTENNA GAINS AND LINE LOSSES
(ALL GAIN VALUES IN DECIBELS RELATIVE TO AN ISOTROPIC ANTENNA)

a) Transmitting Antennas Toward Receiving Site

	Corner Reflector Cheyenne Mt. Summit	Rhombic Cp. Carson	Yagi Pikes Peak
Kendrick	10.0	12.5	8.6
Karval	9.9	5.5	8.1
Haswell	10.0	16.8	8.6
Garden City	10.0	18.2	8.6
Anthony	10.0	17.9	8.5
Fayetteville	10.0	17.6	8.5

b) Receiving Antennas other than Dipoles Toward Transmitting Site

	Rhombic Anthony	Rhombic Fayetteville
Cheyenne Mt. Summit	16.4	19.6
Camp Carson	16.2	20.0
Pikes Peak	16.4	19.4

c) Line Losses	Line Loss, decibels
Transmitting Installations	
Cheyenne Mountain	0.1
Camp Carson	1.1
Pikes Peak	1.1
Receiving Installations	
Kendrick	0.7
Karval	0.1
Haswell	0.1
Garden City	0.7
Anthony	2.5
Fayetteville	2.0

Dipole receiving antennas were used at Kendrick, Karval, Haswell, and Garden City. The theoretical gain value used for these antennas was 2.15 db relative to an isotropic radiator.

Line losses were determined either by direct measurements, or by assumption of a loss figure per unit length in accordance with the manufacturer's specifications.

CALIBRATION AND DATA RECORDING

Transmitter output power was determined by direct measurements with radio frequency wattmeters. Complete logs for all transmitting operations were kept. The transmitter output power was held closely to the nominal values; any departure amounting to more than ± 0.1 db was taken into account in the data analysis. Operating frequency was kept within ± 200 cycles, any greater variations were corrected when detected.

All receiver calibrations were made using signal generators which were checked against a similar unit kept in the laboratory and used as a comparison standard only. At all times, suitable 50 ohm pads were inserted between the signal generators and the receivers.

At Kendrick, Karval, and Haswell the output of the receivers was fed into Esterline-Angus graphic recording ammeters which were operated at a chart speed of three inches per hour. Calibrations were accomplished by connecting the signal generator to the receiver input in place of the antenna lead, and marking a series of signal levels on the charts.

At Garden City, Anthony, and Fayetteville, narrow band receivers were used in conjunction with Esterline-Angus recorders as well as with time totalizers similar to the apparatus described in detail by R. P. Decker.[2] The National Bureau of Standards time totalizer essentially contains 10 channels, each of which includes a dc amplifier, and a bistable multivibrator actuating a fast-acting relay. If the receiver output voltage exceeds a preset level on any of the ten channels, the relay is actuated and a motor started which drives a revolution counter until the receiver output voltage drops again below the preset level. All ten channels are preset to the calibration levels which are also marked on the recording charts, and the face of the counter panel is photographed automatically at hourly intervals by a relay-operated camera. Each film frame obtained by this method indicates also the date, the time of the day, and the preset calibration levels.

At Anthony and Fayetteville the recording charts were operated at speeds ranging up to 1.5 inches per minute in order to obtain data suitable for analysis in case of failure of the totalizing or photographic equipment, and in order to provide records for subsequent investigations of fading range and fading rate.

Procedure Used in Analyzing and Presenting Data

In order to utilize the concept of transmission loss it was found convenient to calibrate the signal generator output in terms of decibels below 2 kw transmitter output. This permits direct tabulation of the recorded data in decibels of transmission loss, requiring only corrections for variations in the signal generator calibrations, for transmission line losses, and for transmitter power values different from 2 kw (applicable to the Camp Carson and Pikes Peak Data).

Where time totalizers were used, the differences between hourly counter readings were expressed as percentages of total times between readings, for each level and these were then graphed on probability paper and the median values obtained. Where time totalizers were not available, the Esterline-Angus charts were scaled for the same type of hourly data.

After all hourly medians were determined, they were corrected as noted above, and the resulting hourly medians of transmission loss in db were classified into four groups as follows (for each of the six receiving sites):

Cheyenne Mountain Transmissions	Aug. 6–13
Camp Carson Transmissions	Aug. 6–13
Cheyenne Mountain Transmissions	Aug. 14–20
Pikes Peak Transmissions	Aug. 14–20

[2] R. P. Decker, "Notes on the analysis of radio propagation data," Proc. IRE, vol. 39, pp. 1382–1388; November, 1951.

For each of the final twenty-four groups, a cumulative distribution of all hourly medians was obtained, plotted on probability graph paper, and from each of these the over-all median and the interdecile range of hourly medians (as a measure of the standard deviation of hourly medians) was obtained (see Figs. 4 to 15, pp. 000–000).

Table III shows the over-all median values, expressed as transmission loss in db, obtained for the four transmission periods at all receiving sites. In order to compensate for the difference in the number of recording hours available for each transmission path and the different periods involved, the data were adjusted to the Cheyenne Mountain weighted average in the following manner, considering the fact that the received hourly median fields when expressed in db, are approximately normally distributed.

TABLE III

Transmission Loss Values
Over-all Medians for 100 Mc Measurements
August, 1952

Receiving Site	August 6–13		August 14–20	
	Cheyenne Mountain	Camp Carson	Cheyenne Mountain	Pikes Peak
Kendrick	122.4	137.3	121.3	116.4
Karval	123.9	151.8	123.5	120.2
Haswell	138.2	152.0	138.3	126.5
Garden City	185.0	180.6	185.5	182.0
Anthony	177.8	177.1	181.2	178.6
Fayetteville	194.7	191.3	195.0	193.0

Above data adjusted to Cheyenne Mountain weighted average.

	Cheyenne Mountain	Camp Carson	Pikes Peak
Kendrick	121.8	136.7	116.9
Karval	123.8	151.7	120.5
Haswell	138.2	152.0	126.6
Garden City	185.1	180.7	181.6
Anthony	179.4	178.7	180.4
Fayetteville	194.8	191.2	192.8

For each receiving location the weighted average of the Cheyenne Mountain transmission loss was first determined by multiplying the over-all median for the August 6–13 period by the total number of hours contained in the distribution for that period, adding to that quantity a number obtained similarly for the August 14–20 period, and dividing the total by the total number of hours for both periods. This gives the weighted average of the Cheyenne Mountain transmissions. Subsequently the differences between this average and the over-all median for each of the two periods were determined, and the Camp Carson and the Pikes Peak median corrected by algebraic addition of the difference for the appropriate period. The adjusted data are also shown in Table III.

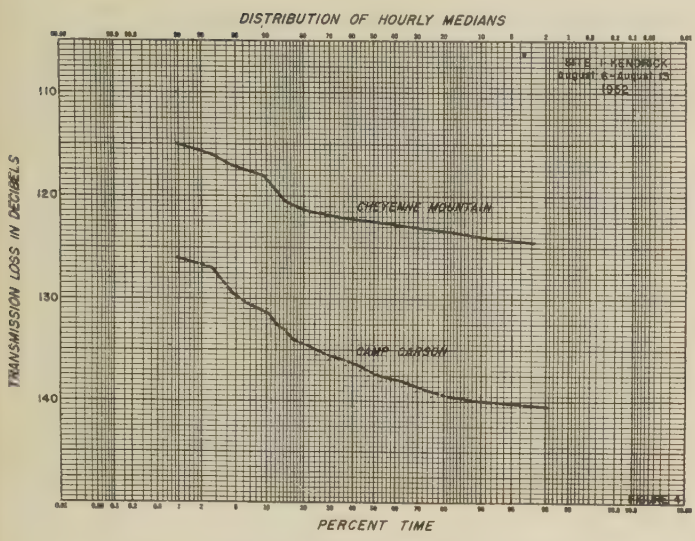


Fig. 4—Site 1, Kendrick, August 6–13, 1952.

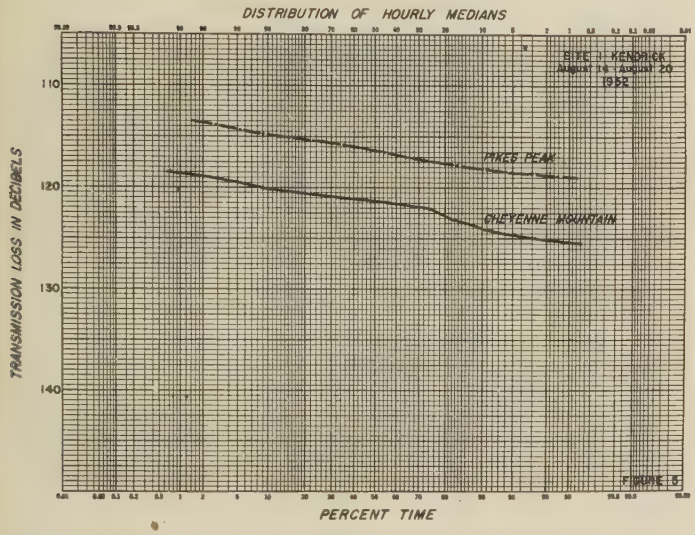


Fig. 5—Site 1, Kendrick, August 14–20, 1952.

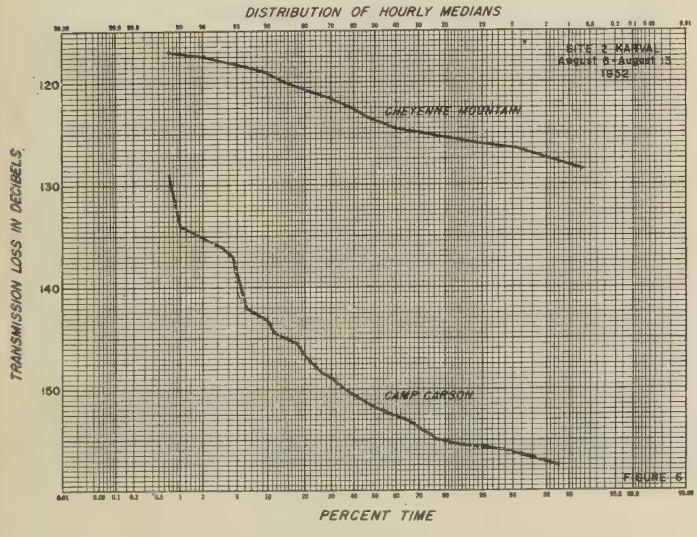


Fig. 6—Site 2, Karval, August 6–13, 1952.

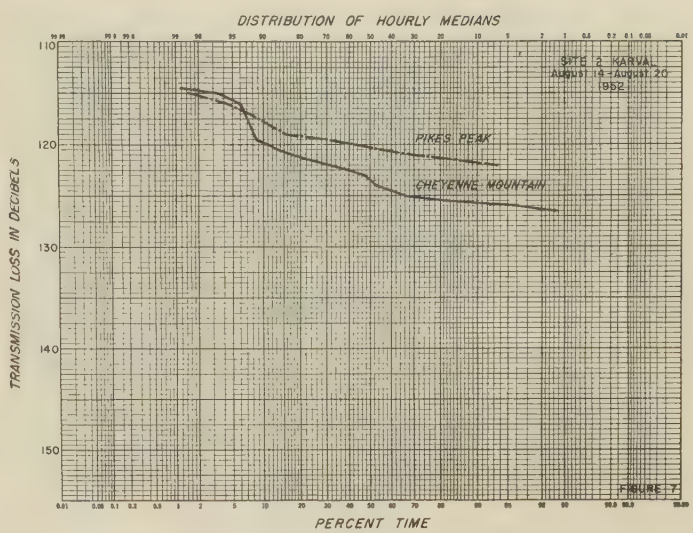


Fig. 7—Site 2, Karval, August 14–20, 1952.

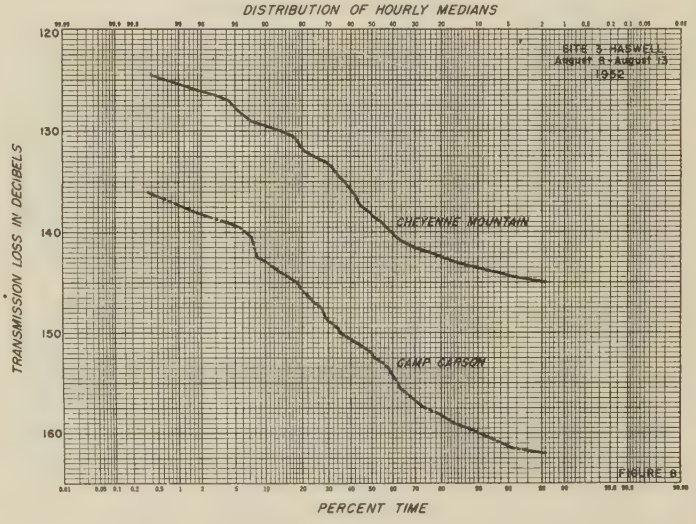


Fig. 8—Site 3, Haswell, August 6–13, 1952.

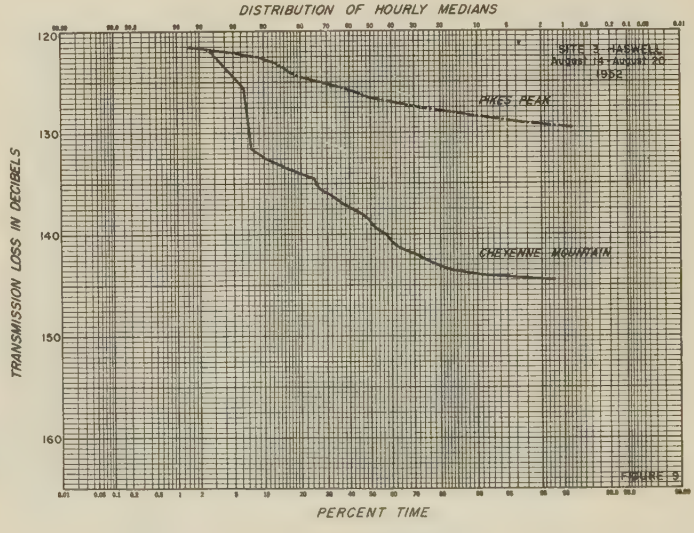


Fig. 9—Site 3, Haswell, August 14–20, 1952.

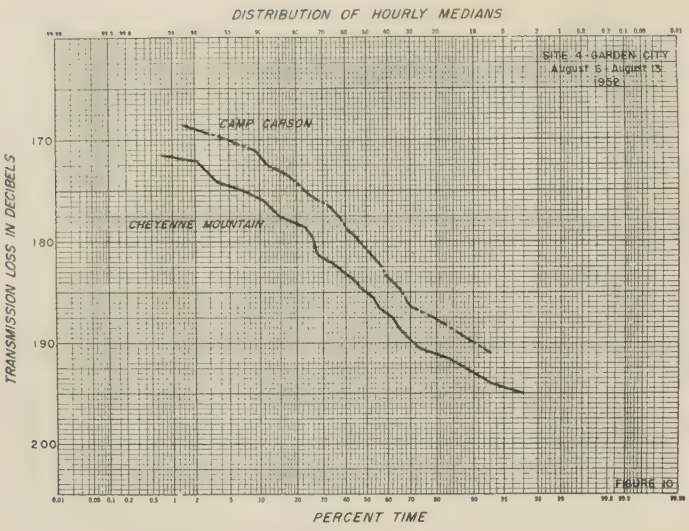


Fig. 10—Site 4, Garden City, August 6–13, 1952.

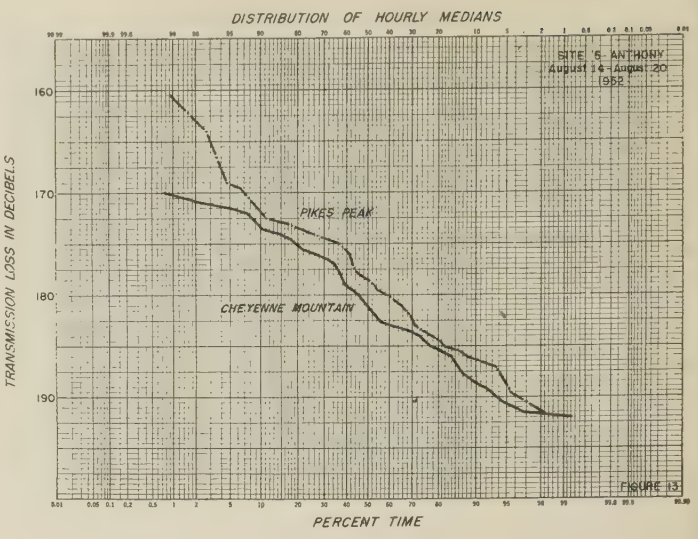


Fig. 13—Site 5, Anthony, August 14–20, 1952.

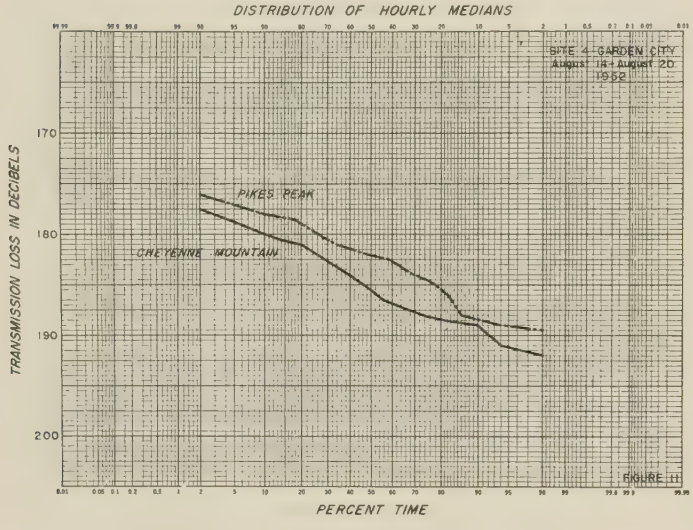


Fig. 11—Site 4, Garden City, August 14–20, 1952.

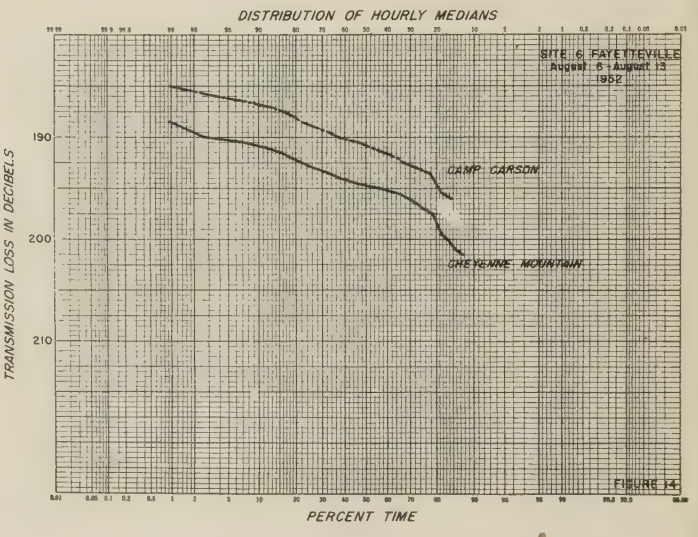


Fig. 14—Site 6, Fayetteville, August 6–13, 1952.

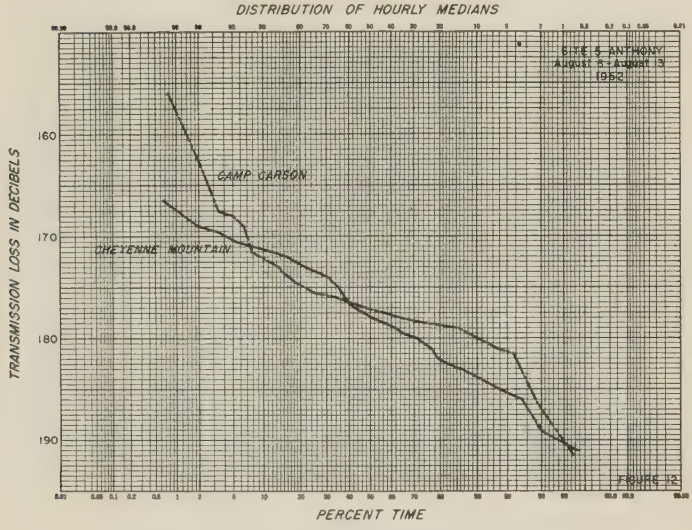


Fig. 12—Site 5, Anthony, August 6–13, 1952.

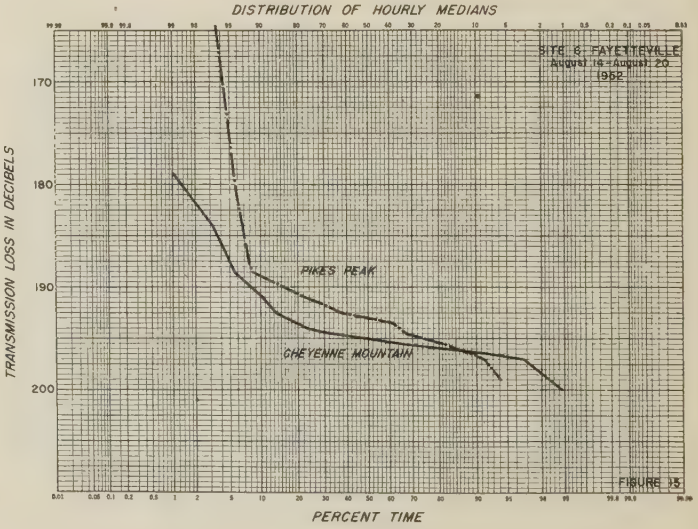


Fig. 15—Site 6, Fayetteville, August 14–20, 1952.

The interdecile range of hourly medians in Table IV was adjusted in a similar way; but in this case the squares of the db values (a measure of the variances), were multiplied by appropriate number of hours instead of first-power values used in adjusting the medians.

TABLE IV

INTERDECILE RANGE
(DIFFERENCE IN DECIBELS OF MEDIANS EXCEEDED 10% AND 90% OF THE TOTAL RECORDED HOURS)
AUGUST, 1952

Receiving Site	August 6–13		August 14–20	
	Cheyenne Mountain	Camp Carson	Cheyenne Mountain	Pikes Peak
Kendrick	5.4	8.6	3.4	3.3
Karval	7.0	12.3	5.9	4.1
Haswell	14.1	17.1	11.5	5.8
Garden City	17.3	17.5	9.0	10.5
Anthony	12.6	7.8	15.2	14.5
Fayetteville	11.5	10.7	5.6	7.7

Above data adjusted to Cheyenne Mountain weighted average.

	Cheyenne Mountain	Camp Carson	Pikes Peak
Kendrick	4.4	7.6	4.3
Karval	6.6	11.9	4.8
Haswell	13.0	14.9	7.3
Garden City	15.5	15.7	17.0
Anthony	13.8	9.0	13.1
Fayetteville	9.2	8.4	11.3

It was felt that the number of recorded hours available was not sufficient to permit a further breakdown of the results in order to determine diurnal variations. The total number of hours for which useful recorded data were obtained is shown in Table V.

TABLE V

NUMBER OF USEFUL RECORDING HOURS FOR COMPARISON RUNS
AUGUST, 1952

Receiving Site	August 6–13		August 14–20	
	Cheyenne Mountain	Camp Carson	Cheyenne Mountain	Pikes Peak
Kendrick	53	53	72	65
Karval	72	72	40	33
Haswell	83	87	72	65
Garden City	67	71	25	25
Anthony	75	70	63	54
Fayetteville	52	52	47	42

Finally, all transmission loss values for the adjusted medians were converted into decibels of attenuation relative to the free space field. This was done in accordance with methods described previously.[3] It should be noted that this conversion makes use of the antenna gains determined by assumptions or close-in measurements, and for the more distant receiving sites does not take into account the "loss-in-gain" concept which appears as one of the conclusions of the Scattering Theory, as applied to long distance tropospheric propagation.[4]

Table VI shows the adjusted results in db relative to the free space field compared to computed values for the sites within the radio horizon and in the diffraction zone (Kendrick, Karval, and Haswell). The computed values are based on the assumption of an effective radius of the earth corresponding to normal refraction, which is applicable to the paths concerned in accordance with available meteorological data,[5] and the approximation of the actual path profiles by second degree curves fitted to appropriate portions of the paths.[6,7]

TABLE VI

COMPUTED AND MEASURED FIELDS IN DECIBELS BELOW FREE SPACE VALUES
AUGUST, 1952

Measured values adjusted to average of Cheyenne Mountain runs.

	Cheyenne Mountain		Camp Carson		Pikes Peak	
	Computed	Measured	Computed	Measured	Computed	Measured
Kendrick	17.6	23.5	31.5	41.4	6.1	15.5
Karval	9.7	22.3	37.5	46.1	−5.0	16.1
Haswell	28.4	34.0	57.5	54.9	21.0	20.1
Garden City	—	73.5	—	77.5	—	68.2
Anthony	—	77.3	—	84.4	—	76.5
Fayetteville	—	91.9	—	96.4	—	88.1

The calculation of expected fields for receiving locations far beyond the radio horizon is treated in a paper delivered at the 1953 IRE National Convention.[4]

The adjusted median values from Table 6 are shown on Fig. 16, plotted vs distance. For greater clarity in presentation a logarithmic distance scale has been chosen. The lines connecting various points denoting results for the receiving sites should not be construed as representing conditions at distances between the actual locations. Similarly, the line indicating the radio horizon is only approximate, and serves to illustrate the location of the sites relative to the radio horizon.

In a similar way the adjusted interdecile range is plotted vs distance and shown in Fig. 17 illustrating the expected signal variations as a function of distance.

DISCUSSION OF RESULTS

Before attempting discussion or evaluation of these measurements, the following should be kept in mind:

[3] K. A. Norton, "Transmission loss in radio propagation," PROC. IRE, vol. 41, pp. 146–152; January, 1953.

[4] J. W. Herbstreit, K. A. Norton, P. L. Rice, and G. E. Schafer, "Radio wave scattering in tropospheric propagation," 1953 IRE CONVENTION RECORD, *Pt. 2, Antennas and Communication*, pp. 85–93.

[5] B. R. Bean, "The geographical and height distribution of the gradient of refractive index," PROC. IRE, vol. 41, pp. 549–550; April, 1953.

[6] K. A. Norton, "Transmission loss of space waves propagated over irregular terrain," TRANS. IRE, vol. AP-3, pp. 152–166; August, 1952.

[7] K. A. Norton, "Calculation of ground wave field intensity over a finitely conducting spherical earth," PROC. IRE, vol. 29, pp. 623–639; December, 1941.

a) The calibration check of the signal generators could be made only after transportation over considerable portions of rough roads. It is unknown in what way or to what extent the signal generators may have been affected by transportation.
b) Although considerable effort was made to match antennas and receiver inputs to the nominal 50-ohm impedance, the stability of these values is not certain. Any mismatch in the antenna-receiver chain, or in the generator-receiver chain would affect all calibrations.
c) The experiments were conducted over a limited number of hours, and, due to some equipment failures, not all hours of transmission could be utilized at all receiving sites. The method of adjusting all data to the weighted average of the Cheyenne Mountain transmissions is not necessarily ideal, but constitutes a reasonable approach to the correct values.

Altogether it is felt that the relative values of the interdecile range could be determined more accurately than the absolute values of transmission loss. Apart from calibration difficulties, the latter are also subject to variation in antenna gains from the values assumed. It is very apparent that there is not much correlation between computed within line-of-sight and diffraction fields, and the actual measurements. This may be chiefly attributed to the oversimplification one has to resort to in order to represent the rough terrain by a second-degree surface. The chief factor contributing to the computed field seems to be the character of the terrain in the vicinity of the receiving antenna, where, within the radio horizon, reflections are thought to take place. Studies are now under way to obtain sufficient data on transmissions within line-of-sight over various paths at various frequencies to correlate departures from computed values with relative terrain roughness and other features.

However, the results obtained from the August measurements definitely confirm the existence of relatively high fields far beyond the radio horizon. Beyond a distance of approximately 100 to 200 miles from the radio horizon the median field does not decrease at the high rate of the diffraction region, and the deviations from the medians tend to decrease from their maximum values at approximately 100 miles beyond the radio horizon (see Figs. 16 and 17). In interpreting Figs. 16 and 17, one should also bear in mind that the lines connecting the points which represent measurements at a particular site are drawn only to show the trend for the three transmitting heights compared. Especially a measuring location between Haswell and Garden City (at about 150 miles from the transmitters) would have been helpful in establishing the distance at which maximum long-term signal variations occur.

The graph shows clearly that the influence of the antenna height on the over-all medians as well as on expected signal variations diminishes with distance from the transmitter.

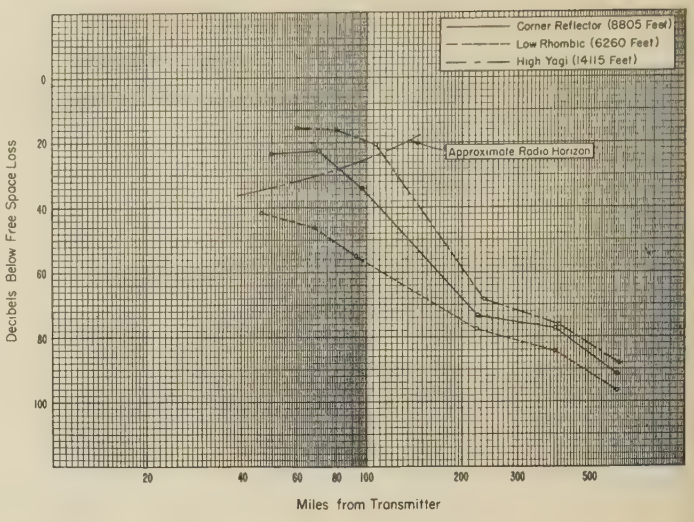


Fig. 16—Transmission loss relative to free space, 100 mc, August, 1952 (assuming antenna gains are realized).

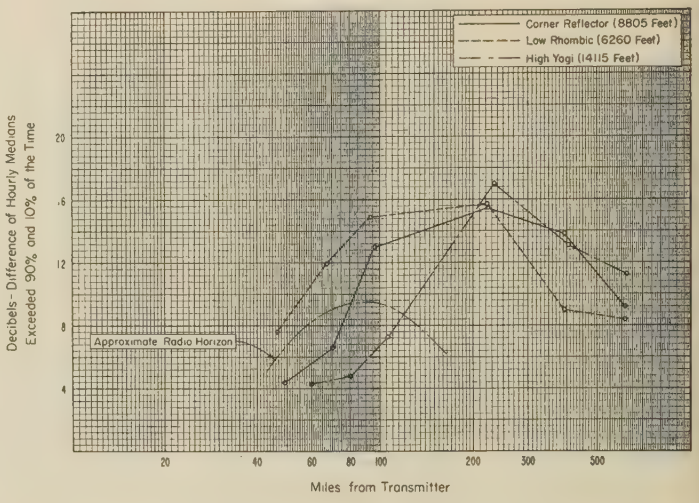


Fig. 17—Variations of hourly medians, 100 mc, August, 1952, for three transmitting antennas.

Conclusion and Recommendations

It is concluded that the measurements presented in this report are useful for the study of long distance tropospheric propagation of very high frequency signals emitted from various types of antennas located at various elevations. They should be supplemented by:

a) Identical experiments at other times of the year.
b) Similar experiments at other frequencies.
c) Study of short-term variations and the fading rate.

It is also desirable to increase the reliability of the data by better methods of calibration of the equipment, and comparison with even more reliable standards of reference. The procedure for antenna gain and pattern measurements may also be improved. It is expected that work now under way on application of Booker-Gordon scattering theory to propagation studies will offer a more realistic approach to effective gain figures of antennas operating on long path transmissions.[4]

The Effect of Superrefractive Layers on 50–5,000 MC Nonoptical Fields*

E. E. GOSSARD† AND L. J. ANDERSON‡

Summary—**A body of radio data taken over an 80 nautical mile overwater link on four frequencies is analyzed, and the relationships to the vertical structure of the atmosphere determined statistically. The results are compared with the attenuation coefficients yielded by the theory for the first mode of a bilinear model as given by Furry.**

It is found that observation agrees reasonably well with theory at about 100 mc, but departs considerably for lower or higher frequencies. It is found that for a given height of the top of the superrefractive layer, an optimum frequency generally exists for which attenuation is a minimum.

IT HAS BEEN recognized for some time that no existing method based solely on theory is adequate for predicting radio fields at all tropospheric frequencies in the diffraction region beyond the optical horizon. In order to predict such fields from meteorological data, it is thus necessary to rely upon empirical methods.

This paper describes a statistical investigation of the relationship of observed radio fields to certain critical parameters describing the meteorological profiles. The empirical results thus obtained are compared with the results of mode theory. An attempt has been made to keep the method as objective as possible and reduce assumptions to a minimum.

The Data

The radio data used were taken over an 82 nautical mile link between San Diego and San Pedro (Fig. 1). The transmitters and receivers were located 100 feet above sea level and the path was entirely over water.[1] Four frequencies were used: 52, 100, 547, and 5,000 mc. The fields were recorded in db above 1 μv and the respective free space fields were 64, 56, 38, and 68 db above 1 μv. Although detailed wiresonde data were available for some of the time during which the link was in operation, it was decided to use radiosonde data in this analysis since the results would then be directly applicable to other links for which only standard radiosonde data were available. In the Pacific Standard Time zone radiosonde ascents are made starting at 7 A.M. and at 7 P.M., hence the radio fields tabulated were averages of the hourly maximum fields recorded between 6 and 9 o'clock, both A.M. and P.M.

Analysis of the Data

All available San Diego radiosonde observations taken while the links were in operation were plotted in terms of a modified refractive index, B, where B is defined by

$$B = (n - 1) \times 10^6 + .012h.$$

n is refractive index and h is height in feet. Those cases for which the refractive index profile was relatively simple in form were then selected for the analysis. Specifically, those cases were rejected which required more than a trilinear model to adequately represent the sounding. Thus, only one superrefractive layer—either surface based or elevated—was examined (Fig. 2).

To determine the form of regression equation to be used in the analysis, scatter diagrams of all data for the various frequencies were plotted to determine the general form of the dependence of field strength on height and intensity of the layer. A previous study by Hoffmann (unpublished) had shown that layer thickness Δh was not statistically significant. The dependence of the attenuation coefficient on the character of the layer according to mode theory was also taken into account in

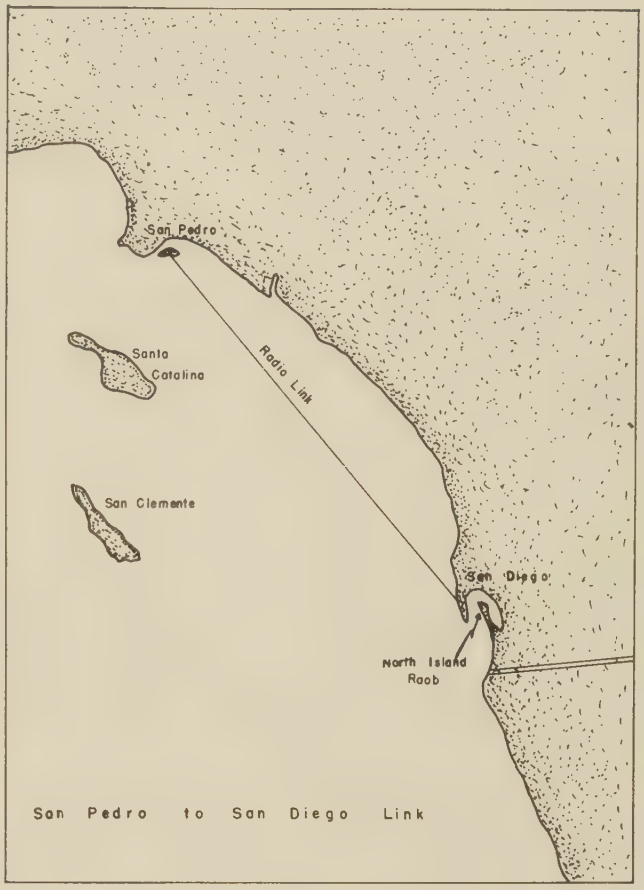


Fig. 1—Map showing location of links.

* Manuscript received by the PGAP, July 13, 1955.
† U. S. Navy Electronics Lab., San Diego 52, Calif.
‡ Smyth Research Associates, San Diego.
[1] J. B. Smyth and L. G. Trolese, "Propagation of radio waves in the lower troposphere," PROC. IRE, vol. 35, pp. 1198–1202; November, 1947.

TABLE I

Freq mcs	Class*	N	r_{12}	r_{13}	$r_{12.3}$	$r_{13.2}$	$r_{1(23)}$	$s_{1.2}$	$s_{1.3}$	$s_{1.23}$	b	c
52	E	104	0.522	−0.056	0.523	−0.075	0.521	6.10	7.14	6.10	25.6	− 3.66
52	S.B.	58	0.622	−0.035	0.685	−0.355	0.686	5.79	7.42	5.41	25.7	− 8.1
100	E	111	0.543	−0.445	0.644	−0.578	0.729	6.07	6.47	4.95	30.7	−26.8
100	S.B.	72	0.743	−0.419	0.742	−0.416	0.793	7.17	9.71	6.51	41.2	−21.1
547	E	114	0.416	−0.524	0.502	−0.593	0.679	10.9	10.3	8.77	40.1	−48.2
547	S.B.	93	0.210	−0.698	0.713	−0.858	0.865	14.1	10.3	7.24	45.9	−50.4
5,000	T	78	0.282	−0.543	0.289	−0.545	0.596	15.5	13.6	13.0	16.3	−45.2

*E = Elevated cases, S.B. = Surface Based cases, T = Total cases.

selecting the regression equation. The final equation used was

$$F = a + b \log \Delta B + c \log h$$

where F is the radio field strength in db below the free space field. ΔB is the refractive index at the surface minus that at the height h (the top of the superrefractive layer) (Fig. 2).

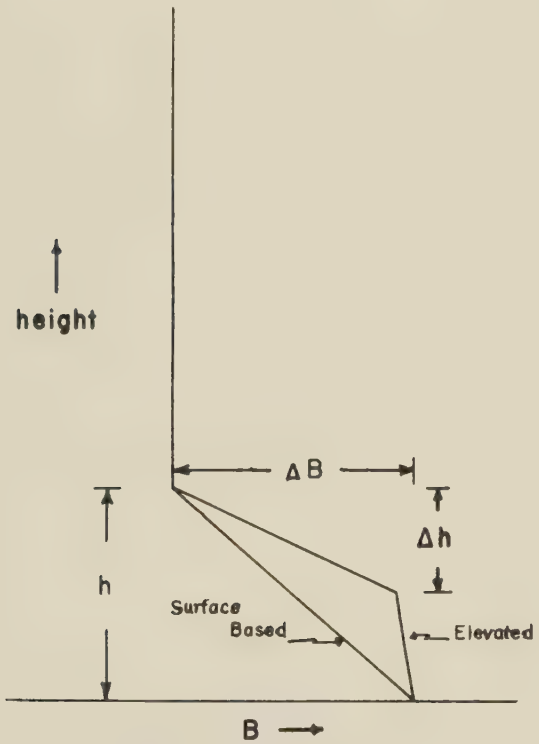


Fig. 2—Simplified trilinear and bilinear B curves illustrating h, ΔB and Δh.

For the three lower frequencies sufficient data were available to permit the sample to be divided into two classes:

a) superrefractive layer surface based.
b) superrefractive layer elevated.

Results of Analysis

The results of the statistical analysis are listed in Table I, above where:

N = number of observations in sample
r_{12} = total correlation between signal and log ΔB
r_{13} = total correlation between signal and log h
$r_{12.3}$ = partial correlation between signal and log ΔB
$r_{13.2}$ = partial correlation between signal and log h
$r_{1(23)}$ = multiple correlation of signal with log ΔB and log h
$s_{1.2}$ = residual standard deviation (effect of log ΔB removed)
$s_{1.3}$ = residual standard deviation (effect of log h removed)
$s_{1.23}$ = residual standard deviation (effect of log ΔB and log h removed)
b = regression coefficient for log ΔB
c = regression coefficient for log h

A covariance analysis of the data indicated the difference in the regression coefficient between the surface based and elevated layers was not statistically significant at the 0.05[2] significance level, although the difference in the 100 mc "b" coefficients was almost significant. Therefore, for prediction purposes it is not worthwhile to separate the two classes and the weighted average of the coefficients was determined.

The multiple and partial correlation coefficients were in all cases significant at the 0.05 level, often at the 0.01 level and sometimes at the 0.001 level.

The plot of the regression coefficients vs frequency is shown in Fig. 3.

Discussion

The fact that there is no significant difference between the surface based and elevated cases makes it reasonable to compare the results of the analysis with those predicted by bilinear mode theory. In this model the critical meteorological parameters may be expressed in terms of h and ΔB, and the solution for the first mode is presented as a function of height (the height gain function) and a function of distance involving an attenuation coefficient.[3] An effort was therefore made to obtain attenuation coefficients from the statistical anal-

[2] If correlation is significant at the 0.05 level one chance in 20 exists that the observed relation is due to chance.

[3] W. H. Furry, "Theory of characteristic functions in problems of anomalous propagation," *M.I.T. Rad. Lab. Report 680*; February, 1945.

ysis in order to compare them with those of theory. In order to keep the results as empirical and objective as possible, it was decided to approximate the effect of the height gain function by simply determining the field of the radio horizon by some standard method such as that outlined by Domb and Pryce.[4] The attenuation coefficients then represent simply the linear rate at which the field (in db) decreases with distance beyond the horizon

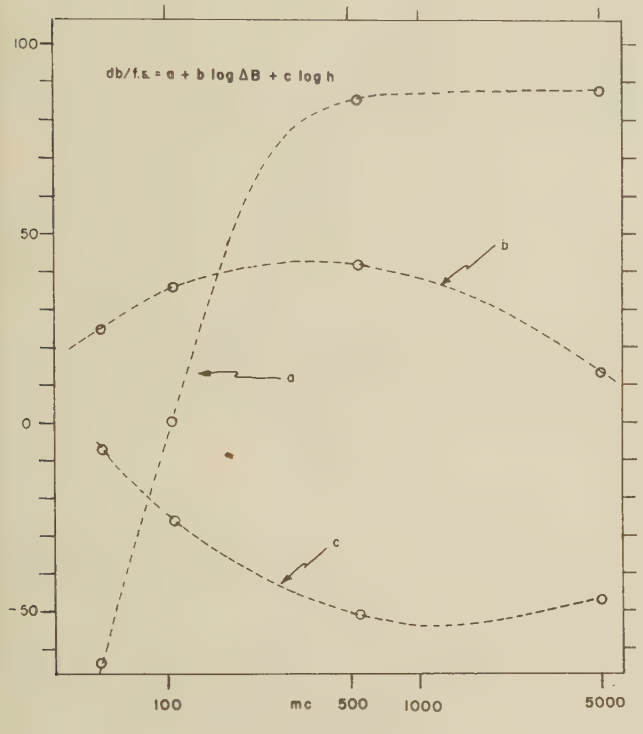


Fig. 3—Plot of regression coefficients vs frequency for San Pedro-San Diego links.

point (Fig. 4). The effect of superrefraction on the horizon field can be accounted for approximately by altering the effective radius of curvature of the earth, thus moving the horizon point farther from the transmitter and lowering the field. This method breaks down as trapping is approached, yielding negative attenuation coefficients. For large ΔB's and thin layers, a further difficulty is encountered in assigning an effective earth's radius. If we assume, however, that the height gain function corresponding to the first mode of the bilinear model is valid for such layers, we can solve for an effective attenuation coefficient needed to yield the observed fields. Actually it developed that the attenuation coefficients obtained by this approach were similar even for these extreme conditions to those obtained assuming a horizon field that was independent of refraction. This must be considered accidental, since the horizon field under large ΔB and thin layer conditions must depart considerably from the standard case in a way that cannot be taken into account by any simple procedure, but the similarity of this result with that obtained by using the height gain function adds confidence in the *fixed horizon point* technique—particularly when the superrefraction is not extreme.

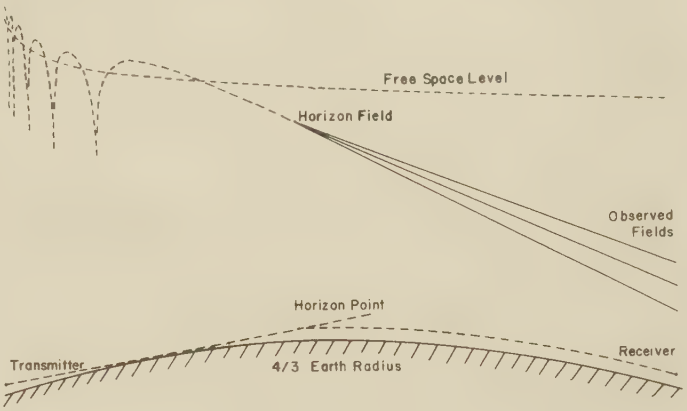


Fig. 4—Curved earth representation of horizon ray, showing diffraction region where attenuation coefficients are determined. Solid lines show field strength vs range for various attenuations.

The empirical attenuation coefficients so obtained are compared with the theoretical in Figs. 5–8.

Obviously, the regression analysis only yields values above the *knee* in the attenuation curves, and only data above this critical h were examined. Furthermore, the noise level of the receiver was above the standard diffracted field, so the data on which this analysis is based were all taken under somewhat superrefractive conditions. To properly weight the *near standard* area of the graphs, the fields were assumed to go to the standard diffracted field as $\Delta B \rightarrow 0$ and the curves in this region obtained by interpolation.

Interesting features of the analysis are the following:

1) In the present analysis the difference in the regression coefficients between surface based and elevated cases was not significant.

2) The departure of the observations from mode theory is such that fairly good agreement is noted at about 100 mc, becoming worse as frequency is either increased or decreased.

3) The effect of elevating the layers does not diminish the field nearly as rapidly as would be anticipated by theory.

4) For a given layer height, the variation of the regression coefficients with frequency as shown in Fig. 3 leads to an optimum frequency for which the field is maximum (minimum attenuation coefficient). As the layer gets higher, this maximum field occurs at a lower and lower frequency. This is principally the result of the change in slope of the attenuation lines with increasing frequency. Physically it could be explained in terms of two opposing trends resulting from scattering within the superrefractive layer. Such scattering (perhaps by gravity waves) has been suggested by Anderson and

[4] C. Domb and M. H. L. Pryce, "The calculation of field strengths over a spherical earth," *JIEE*, vol. 99, pp. 325–339; September, 1947.

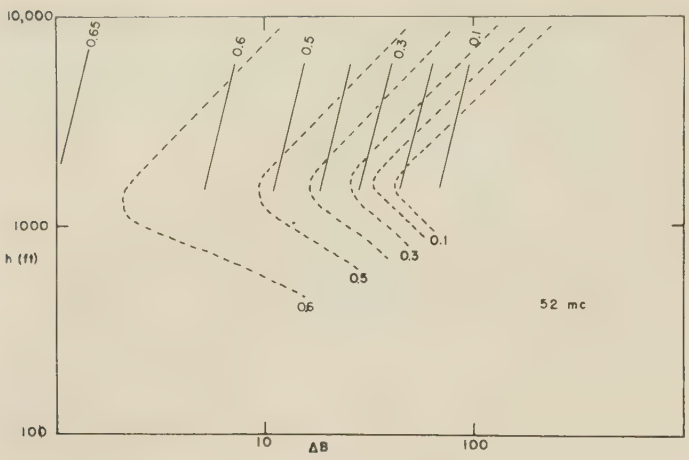


Fig. 5—Attenuation coefficients 52 mc. Dashed curves bilinear mode theory.

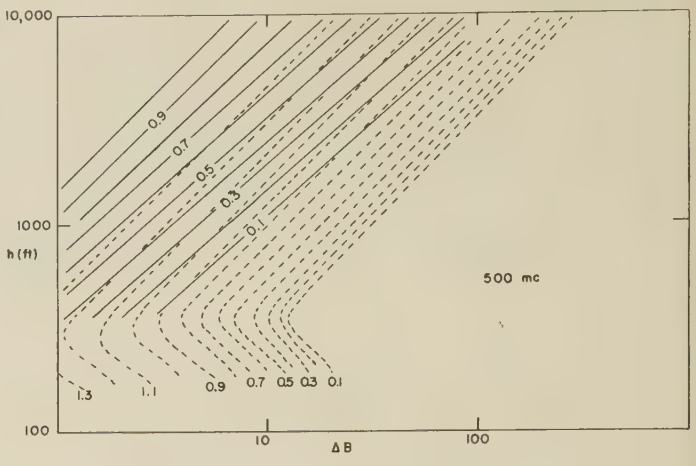


Fig. 7—Attenuation coefficients 547 mc. Dashed curves bilinear bode theory.

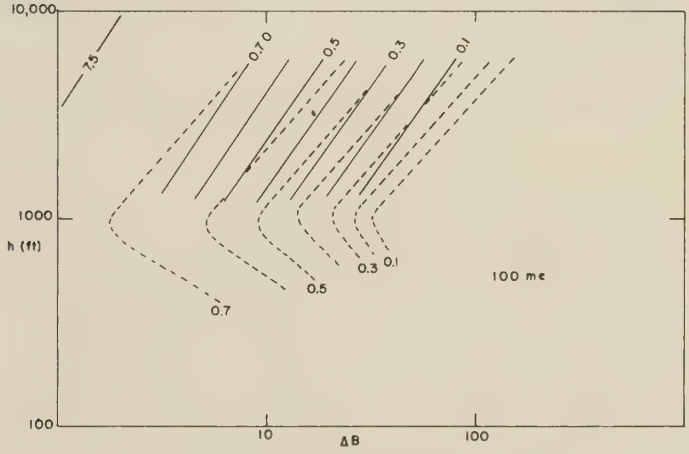


Fig. 6—Attenuation coefficients 100 mc. Dashed curves bilinear mode theory.

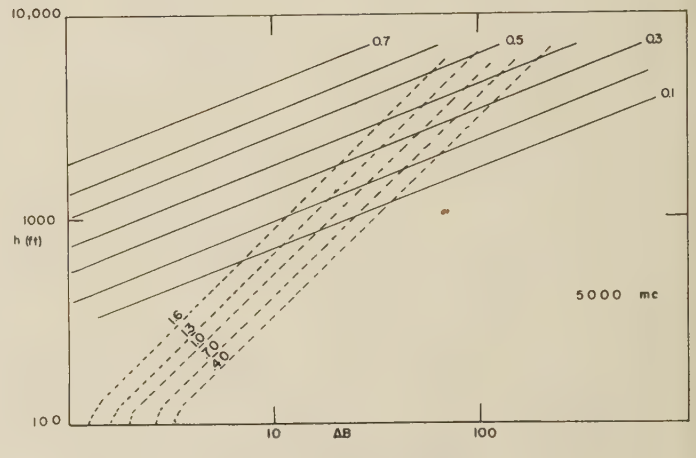


Fig. 8—Attenuation coefficients 5,000 mc. Dashed curves bilinear mode theory.

Smyth.[5] Thus, as the frequency increases scattering becomes *more* effective, but for a given height of the layer increasing the frequency will *lessen* the over-all effect of the layer. Therefore if the layer is very high the field would decrease with increasing frequency even though the scatterers are more effective. For h between 2,000 and 3,000 feet the optimum frequency is about 400 mc.

Application

Although the purpose of this paper is to present the results of a statistical analysis of experimental data, it should be pointed out that the basic purpose in doing the work was to develop a method of predicting radio field strength beyond the horizon. The purpose in using only radiosonde data for refractive index profiles and in presenting the results as attenuation coefficients beyond the horizon point, was aimed at the development of an objective method which could be applied to links of differing geometry as well as differing climatic conditions. Work is proceeding on the application of the method to such other links with the purpose of comparing predicted fields with observation. The results of these comparisons will be presented in a subsequent paper.

Acknowledgment

The authors are grateful to Dr. J. B. Smyth for guidance and for contributing some of the basic ideas contained in this paper, to Dr. W. C. Hoffman for statistical guidance and to F. A. Sabransky for much of the data processing.

[5] L. J. Anderson and J. B. Smyth, "The effect of uniform layers on the propagation of radio waves," Trans. IRE, vol. AP-2, pp. 28–34; March, 1952.

communications

Effect of the Ground Screen on the Field Radiated from a Monopole*

J. R. WAIT†

SOME INTEREST has been shown recently in the radiation characteristics of an L.F. vertical antenna with a radial wire ground system. The investigations[1–3] have been primarily concerned with the input impedance at the terminals of the antenna. This was shown to be mainly a function of the number and length of radials and the ground conductivity. It was assumed in most of this work that the ground wave field for a given current on the antenna was not appreciably affected by changes in size of the ground screen. Under this assumption, the radiation efficiency of the antenna is determined mainly by the input resistance.

In an earlier paper,[2] an approximate method was given which was suitable for estimating the dependence of the ground wave field on the size of the ground screen. It is the purpose of this note to show that quantitative results can be obtained which support our earlier contention that the ground screen has only a small effect on the ground wave field intensity for a specified current on the antenna.

The ground screen is assumed to be a perfectly conducting disc of radius a lying on a homogeneous flat ground of conductivity σ. Choosing a cylindrical coordinate system (ρ, ϕ, z) the antenna is considered to be coincident with the positive z axis and the surface of the ground is defined by $z=0$. Denoting $H_\phi(\rho, z)$ as the magnetic field of the antenna in the absence of any ground screen and $\Delta H_\phi(\rho, z)$ as the change due to the presence of the ground screen, it follows from an earlier paper,[2] that

$$\frac{\Delta H_\phi(\rho, z)}{H_\phi(\rho, z)} \simeq -\frac{\beta\eta}{\eta_0}\frac{e^{-i\beta[\rho^2+z^2]^{1/2}}}{[\rho^2+z^2]^{1/2}} \cdot \int_{\rho'=0}^{a} \frac{H_\phi^\infty(\rho', 0)}{H_\phi^\infty(\rho, z)} J_1(\beta\rho' \cos\theta)\rho' d\rho' \qquad (1)$$

where $\beta = 2\pi/\lambda$, λ = free space wavelength, $\eta = (\mu\omega/\sigma)^{1/2} e^{i\pi/4}$ (surface impedance of the ground), $\mu = 4\pi \times 10^{-7}$, ω = angular frequency, $\eta_0 = 120\pi$ (intrinsic impedance of free space), $\theta = \tan^{-1} z/\rho$, and where J_1 is the Bessel function of the first type. In the above, H_ϕ^∞ refers to the field of the antenna over a flat perfectly conducting ground. This expression for the fractional change of the magnetic field is approximate and neglects terms of higher order in (η/η_0). It also assumes that the attenuation of the ground wave has a dependence with distance which is independent of the size of the ground screen. $\Delta H_\phi/H_\phi$ can be regarded as the fractional change of the effective height of the antenna due to the presence of the ground screen.

As an example a quarter-wave monopole antenna is considered with an assumed sinusoidal current distribution. In this case

* Manuscript received by the PGAP, December 15, 1955.
† National Bureau of Standards, Boulder, Colo.
[1] F. R. Abbott, "Design of buried R.F. ground systems," PROC. IRE, vol. 40, pp. 846–851; July, 1952.
[2] J. R. Wait and W. A. Pope, "The characteristics of a vertical antenna with a radial conductor ground system," *App. Sci. Research*, vol. B4, pp. 177–195; March, 1954.
[3] J. R. Wait and W. A. Pope, "Input resistance of L. F. unipole aerials," *Wireless Eng.*, vol. 32, pp. 131–138; May, 1955.

$$H_\phi^\infty(\rho', 0) = \frac{-iI}{2\pi\rho'} e^{-i\beta[\rho'^2+(\lambda/4)^2]} \tag{2}$$

and since $\rho \gg \lambda$;

$$H_\phi^\infty(\rho, z) \simeq \frac{-iI}{2\pi\rho} e^{-i\beta[\rho^2+z^2]^{1/2} \cos(\pi/2 \sin\theta)}. \tag{3}$$

It then follows that the fractional change of the field is given by

$$\frac{\Delta H_\phi(\rho, z)}{H_\phi(\rho, z)} = \frac{-\beta\eta}{\eta_0} \frac{\cos\theta}{\cos\left(\frac{\pi}{2}\sin\theta\right)} \cdot \int_0^a e^{-i\beta[\rho'^2+(\lambda/4)^2]^{1/2}} J_1(\beta\rho' \cos\theta) d\rho'. \tag{4}$$

The integral, apparently, cannot be evaluated in closed form. It is not too difficult, however, to evaluate it by a graphical method. This has been done for the case $z=0$ which corresponds to the ground wave field. Letting

$$\frac{\Delta H_\phi(\rho, 0)}{H_\phi(\rho, 0)} = \delta[X_1 + iX_2] \tag{5}$$

with $\delta = |\eta/\eta_0|$, the integrals to consider are

$$X_1 = -\int_0^{2\pi a/\lambda} \cos\left[(p^2 + \pi^2/4)^{1/2} - \pi/4\right] J_1(p) dp \tag{6}$$

and

$$X_2 = \int_0^{2\pi a/\lambda} \sin\left[(p^2 + \pi^2/4)^{1/2} - \pi/4\right] J_1(p) dp. \tag{7}$$

Some numerical values of X_1 and X_2 are given in the following Table I.

TABLE I

$2\pi a/\lambda$	X_1	X_2
0.0	0.000	0.000
0.5	−0.042	0.040
1.0	−0.130	0.181
1.5	−0.211	0.417
2.0	−0.209	0.700
2.5	−0.102	0.947
3.0	0.042	1.093
3.5	0.155	1.131
4.0	0.171	1.133
4.5	0.113	1.178
5.0	0.050	1.300
5.5	0.050	1.468
6.0	0.119	1.612
6.5	0.205	1.674

The ratio of the field with the ground screen to the field $H_\phi(\rho, z)$ without the ground screen is then given by

$$1 + \frac{\Delta H_\phi(\rho, z)}{H_\phi(\rho, z)} = Ae^{i\Phi} \tag{8}$$

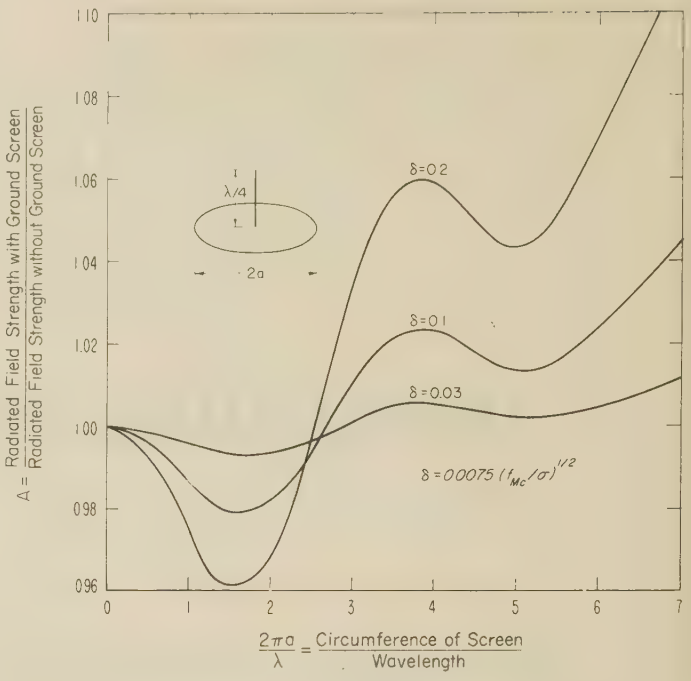


Fig. 1—The ratio of the ground wave field strengths of the antenna with and without the ground screen for a specified antenna current.

where the amplitude ratio A is given by

$$A = [(1 + \delta X_1)^2 + (\delta X_2)^2]^{1/2} \tag{9}$$

and the phase factor is Φ given by

$$\Phi = \tan^{-1} \frac{\delta X_2}{1 + \delta X_1}. \tag{10}$$

Since δ has already been considered small, these are given adequately by

$$A \simeq 1 + \delta X_1$$

and

$$\Phi \simeq \delta X_2 \text{ (radians)}.$$

The factor δ can be obtained conveniently from the following relation

$$\delta = 0.0075\ (f_{mc}/\sigma)^{1/2}$$

where f_{mc} is the frequency in mc and σ is the ground conductivity in mhos/meter. (Since displacement currents in the ground have been neglected throughout, the formulas are valid only when $\delta^2 \ll 1$.) Using (9), the amplitude ratio A is shown plotted in Fig. 1 as a function of $2\pi a/\lambda$, the circumference of the ground screen in wavelengths, for typical values of δ.

It is interesting to note that for small screens the ground wave field strength is actually slightly less than it would be in the absence of the screen. As the screen becomes larger, the amplitude ratio increases somewhat, but is still only a few per cent greater than unity for a less than a wavelength. Thinking in terms of the recipro-

cal situation where the monopole is regarded as a receiving antenna and the transmitter is located at some distant point on the surface of the ground, the increase of A above unity is characteristic of a "recovery" effect. Such a phenomenon occurs in ground wave propagation from land to sea.[4]

Another interesting feature of the curves in Fig. 1 is that the minimum values of A occur approximately where the input resistance of the antenna is a minimum. It can be generally concluded, however, that the dependence of the L.F. ground wave field strength on the size of the ground screen is of minor significance compared to the dependence of the input resistance on the size of the screen. It should be mentioned that Page and Monteath[5] have used a similar method to calculate the radiation pattern of a vertical antenna over an irregular ground plane.

ACKNOWLEDGMENT

I would like to thank W. A. Pope who carefully carried out the graphical integration of X_1 and X_2.

[4] G. Millington, "Ground wave propagation over an inhomogeneous smooth earth," *Proc. IEE*, vol. 96, pp. 53–63; January, 1949.

[5] H. Page and G. D. Monteath, "The vertical radiation patterns of medium-wave broadcasting aerials," *Proc. IEE*, vol. 102, pp. 279–297; May, 1955.

Admittance of Thin Antennas*

GIORGIO BARZILAI†

WITH REFERENCE to my recent paper,[1] I would now like to report some additional calculations. They refer to center driven half-wave antennas.

The results are the following:

$$V_{02} - V_{01} = \frac{60}{\Omega}(1.19 + j0.23) \tag{1}$$

$$P_2 - P_1 = \frac{19}{\Omega^2} + \frac{38}{\Omega} \tag{2}$$

$$V_{01} = 73.13 + j42.54 \tag{3}$$

$$P_1 = 36.57. \tag{4}$$

Result (1) was obtained by introducing the current $I_1(z) - I_0(z)$ into the integral expression for the vector potential. The numerical integration was carried out by Simpson's rule, after dividing the antenna into 16 equal parts.

Result (2) was obtained using the method of moments, after dividing the antenna into 4 equal parts. To compute the moments,

$$\int_0^z f_1(z)dz$$

was expressed in terms of the sine and cosine integral functions.

The final results arrived at are

Ω	$\hat{G}_1$	$\hat{G}_{2,\Omega}$	$\hat{G}_{2,\Psi}$	$G_{2,K}$	G^S
10	10.22	9.69	9.61	9.38	9.77
15	10.22	9.85	9.81	9.60	9.94
20	10.22	9.94	9.92	9.74	10.02

It should be noted that in this case, since the current is prescribed at the input terminals, the resistance can be found immediately from the knowledge of the radiated power. For half-wave antennas, and in general for short antennas where the current can be prescribed at or near the input, both components of the admittance can therefore be calculated with good accuracy. From the values of the conductances recorded above, and from the values of the resistances, the following susceptances were found

Ω	$\hat{B}_1$	$B_{2,\Omega}$	$\hat{B}_{2,\Psi}$	$\hat{B}_{2,K}$	$\hat{B}^S$
10	−5.94	−5.05	−4.90	−4.52	−5.68
15	−5.94	−5.36	−5.29	−5.16	−5.78
20	−5.94	−5.50	−5.46	−5.41	−5.82

Admittances are expressed in millimhos.

* Manuscript received by the PGAP, February 14, 1956.

† Instituto Superiore delle Poste e delle Telecomunicazioni, Fondazione U. Bordoni, Roma.

[1] G. Barzilai, "On the imput conductance of thin antennas," TRANS. IRE, vol. AP-3, no. 1, pp. 29–32; January, 1955.

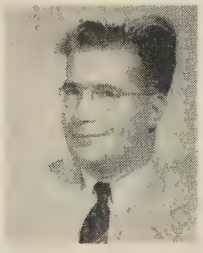
Contributors

L. J. ANDERSON

L. J. Anderson was born in Salt Lake City, Utah, in 1917. He received his A.B. from the University of California at Los Angeles in 1939, and his M.A. from the same university in 1942. He then joined the staff of the U. S. Navy Electronics Laboratory where he was engaged in radar propagation. During his 13 years at NEL, Mr. Anderson also worked on meteorological instrumentation, the development of methods of predicting tropospheric propagation from routine meteorological data, and was later in charge of the radio-meteorology and oceanography work at NEL.

In August, 1955, Mr. Anderson helped in the formation of a private research and consulting firm, Smyth Research Associates, where he is in charge of radio meteorology.

❖



A. P. BARSIS

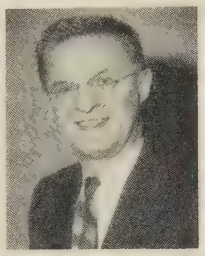
A. P. Barsis (S'48–A'49–M'53) was born in 1916 in Vienna, Austria. He graduated from the Realgymnasium in 1935, and attended the Technische Hochschule in Vienna. He came to the United States in 1939, and from 1940 to 1942, and 1946 to 1952 was employed with George C. Davis, Consulting Radio Engineer, Washington, D. C., where he worked on broadcast allocations and also on the design of directional antennas.

During World War II, Mr. Barsis served as a radio technician with the psychological warfare branch of the Army. He attended George Washington University, and received the B.E.E. degree in 1948. In 1952, he joined the Boulder Laboratories of the National Bureau of Standards, and has been employed in tropospheric propagation research especially in the analysis and evaluation of long-distance tropospheric propagation data provided by the Cheyenne Mountain recording project.

Mr. Barsis is a member of Sigma Tau, an associate member of the Scientific Research Society of America, and is a registered professional engineer.

❖

C. I. Beard was born in Ambridge, Pa., on November 30, 1916. He received the B.S. degree in E.E. from Carnegie Institute of Technology in 1938. One and one-half years at Westinghouse Research Laboratories on dielectrics were followed by two years at M.I.T. as a teaching fellow in physics. In the U. S. Army Signal Corps he served during 1942 and 1943 as an instructor in the Harvard radio and the M.I.T. radar schools. During 1944 and 1945 he was at the Fort Monmouth Signal Corps Publication Agency preparing technical manuals for microwave radar sets. Returning to M.I.T. in 1946, he became a research associate in physics and obtained the Ph.D. degree in physics in February, 1948 performing his research in microwave spectroscopy. In January, 1948 he joined the Magnolia Petroleum Co. Field Research Laboratories and conducted experiments in electromagnetic wave propagation in conducting media. Since August, 1950 he has been at the Applied Physics Laboratory of Johns Hopkins University engaged in microwave spectroscopy and in the propagation of microwave radiation over the surface of the ocean.

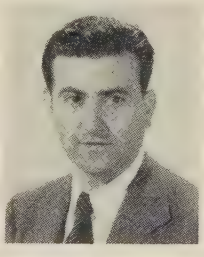


C. I. BEARD

Dr. Beard is a member of the American Physical Society, Tau Beta Pi, and Sigma Xi.

❖

A. D. Berk was born on April 28, 1925, in Istanbul, Turkey, where he received the B.S. degree in E.E. from Robert College in 1947. He continued his studies in the United States at Rensselaer Polytechnic Institute and served as an instructor in the department of electrical engineering of the same institution. After receiving the M.S. degree from R.P.I. in 1951, he continued his studies at the Massachusetts Institute of Technology while serving as a research assistant at the Research Laboratory of Electronics. He received the degree of Doctor of Science from M.I.T. in September, 1954.



A. D. BERK

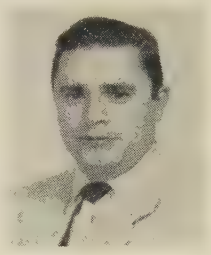
Dr. Berk specializes in problems of electromagnetic theory. Since 1954 he has been a member of the technical staff of the Research Laboratories, Hughes Aircraft Co., Culver City, Calif.

He is a member of Sigma Xi.

❖

For a photograph and biography of E. H. Braun, see page 95 of the January, 1956 issue of TRANSACTIONS OF THE PGAP.

H. N. Chait (A'55) was born in Boston, Mass., on February 15, 1924. He attended Northeastern University, and transferred to Tufts College, obtaining a B.S. degree in electrical engineering in 1945. He has since done graduate work at the University of Maryland.



H. N. CHAIT

After three years service as an electronics technician in the U. S. Navy, Mr. Chait joined the Naval Research Laboratory, doing research and development work in the microwave region. He has done research on the design of signal generators and test equipment. At present he is a member of the Microwave Antennas and Components Branch at NRL doing research on antennas and microwave applications of ferrites.

❖

C. M. Chu was born on March 3, 1919 in Haining, China. He received the B.S. in electrical engineering from National Wu-Han University in 1939, the M.S.E. in 1946 and the Ph.D. in 1952 from the University of Michigan. He was selected as a University Fellow in 1947, 1948, and 1949.

C. M. CHU

Dr. Chu served as an instructor in National Wu-Han University from 1939 to 1941. He worked for the Szechuan Manufacturing Corp. from 1943 to 1945 and for the Laboratory Division of RCA in Princeton, N. J. from 1946 to 1947. In 1951 he became assistant professor of electrical engineering at Detroit Institute of Technology and from 1953 to 1955 was associate professor. He has worked part-time for the Engineering Research Institute of the University of Michigan since 1952 and in 1955 became an associate research engineer. His research work has been in electrical transmission, wave propagation and wave scattering.

He is a member of the Association for Computing Machinery and the Society of Sigma Xi.

❖

S. W. Churchill was born on June 13, 1920 in Imlay City, Mich. He received B.S.E. degrees in both mathematics and chemical engineering from the University of Michigan in 1942. From 1942 until 1946 he worked at the Wood River Refinery of the Shell Oil Co. on the development and production of aviation gasoline and military lu-

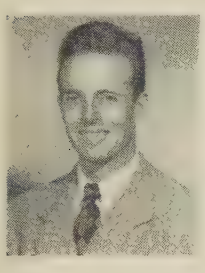
bricants. In 1946 he joined the Frontier Chemical Co. as technical supervisor. He returned to the University of Michigan for graduate work in 1947 and received the Ph.D. degree in 1952.

S. W. Churchill

At the present time Dr. Churchill is associate professor of chemical engineering at the University of Michigan and is active in research and consultation in thermal radiation, convection, chemical kinetics, and combustion. He is a registered professional engineer and is a member of the American Institute of Chemical Engineers, the American Chemical Society, the American Society of Engineering Education, the Combustion Institute, Sigma Xi, and Phi Lambda Upsilon.

❖

E. E. Gossard was born in Eureka, Calif., on January 8, 1923. He received the A.B. degree in meteorology in 1948 and the M.S. degree in physical oceanography in 1951, both from the University of California at Los Angeles.



E. E. Gossard

In 1948 Mr. Gossard joined the U. S. Navy Electronics Laboratory as a meteorologist and has since been engaged in research in the meteorology and micrometeorology of the troposphere and its effect on microwave propagation for the purpose of achieving more accurate operational predictions of radio and radar performance.

He is a member of the American Meteorological Society and American Geophysical Union.

❖

Raymond Justice was born on October 14, 1924 in Loganport, Ind. He received the B.S. degree in 1949 from Purdue University, and the M.A. in 1950 from Ohio State University where he is continuing his graduate studies.

Raymond Justice

Mr. Justice has been employed by the Antenna Laboratory at Ohio State University since 1950. His field of work has been problems related to microwave antennas.

He is a member of Eta Kappa Nu and of Sigma Xi.

❖

M. L. Kales (M'51) was born on August 26, 1910, in New York, N. Y. He received the B.S. and M.S. degrees in mathematics from Massachusetts Institute of Technology in 1933 and 1934, and the Ph.D. in mathematics from Brown University in 1936.

M. L. Kales

From 1936 to 1944 Dr. Kales taught mathematics at Brown University, the University of Michigan, Tulane University, and the University of Maine, successively. In April, 1944, he joined the Radiation Laboratory at M.I.T. At present he is a member of the Microwave Antennas and Components Branch at the Naval Research Laboratory, in Washington, D. C.

Dr. Kales is a member of the American Mathematical Society and Sigma Xi.

❖

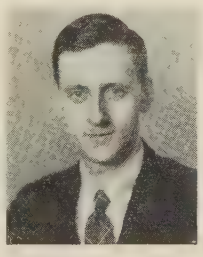
Isadore Katz was born in Philadelphia, Pa., on October 21, 1916. He received the B.S. degree with majors in physics and mathematics from Temple University in 1937. From 1938 until 1941 he worked for the U. S. Weather Bureau. In 1942 he joined the staff of M.I.T. Radiation Laboratory where he was engaged in research in radio-meteorological problems. In 1946 he joined the Naval Research Laboratory staff in Washington, D. C., where he carried out research on microwave propagation and atmospheric physics. Since 1952 he has been with the Applied Physics Laboratory of the Johns Hopkins University where he continues the study of various phases of radar propagation and atmospheric turbulence.

Isadore Katz

He is a member of the American Physical Society and a professional member of the American Meteorological Society.

❖

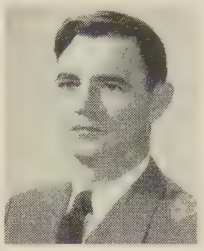
C. A. Levis (S'48–A'52) was born in Berlin, Germany, in 1926. He came to the United States in 1938 and entered Case Institute of Technology in 1944. After two years as an electronic technician in the Navy he returned to Case, receiving the B.S. degree in electrical engineering in 1949. From 1948 to 1949 he was employed as studio engineer at WSRS. He entered Harvard University as Gerard Swope Fellow in 1949 and received the A.M. degree in 1950. Since 1950 he has been a research associate at the Antenna Laboratory of the Ohio State University.



C. A. Levis

He is a member of Tau Beta Pi, Sigma Xi, Eta Kappa Nu, and Pi Mu Epsilon.

R. E. McGavin (A'52) was born in Youngstown, Ohio, on January 13, 1921. He attended Loyola University in Chicago until 1943. After two years in the Navy he completed studies at Niagara University and received the B.A. degree in philosophy in 1946. In 1950 he received the B.A. in physics from the University of Colorado.



R. E. McGavin

He joined the National Bureau of Standards, where he is engaged in the study of tropospheric propagation.

Mr. McGavin is a member of Sigma Pi Sigma and an associate member of the American Institute of Physics and the Research Society of America.

❖

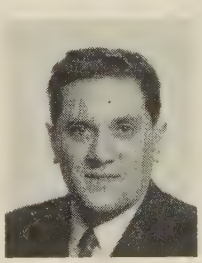
H. E. J. Neugebauer studied mathematics and physics at the University of Berlin from 1924 to 1929. After receiving a graduate degree, he worked with various firms on research and development projects, most of which were connected with optics or colorimetry. In 1935, he received his Doctor's degree from the Technical University of Dresden with a thesis on the theory of multicolor printing which has become important in the development of color correctors by I.P.I. and RCA.

H. E. Neugebauer

Dr. Neugebauer came to Montreal, Canada in 1951 and applied his experience in visual optics to microwaves optics at the Eaton Electronics Research Institute of McGill University. After brief employment as consulting engineer with Adalia Ltd., he joined the newly established Research Laboratories of RCA Victor Co. Ltd., Montreal, as head of the Electro-Magnetics division.

❖

J. A. Ortusi (SM'50) was born in December, 1916 at Poggio di Nazza in Corsica, France. He pursued classical studies at the Marseille Lyceum and the French Polytechnik School and received the Doctor of Science degree in June, 1944.



J. A. Ortusi

Since 1941, he has worked with the Compagnie Generale de TSF in Paris, where he has done research in magnetrons, antennas, TE_0 circular wave, and semiconductors studies. He is now on the technical staff of that organization.

Dr. Ortusi became Laureate of the French Scientific Academy in 1946.

J. H. Richmond (S'50–M'56) was born in Kalispell, Mont., on July 30, 1922. He served in the U. S. Navy as a Chief Electronics Technician from 1940 to 1946 and from 1950 to 1951. In 1950, he graduated summa cum laude from Lafayette College with the B.S. degree in electrical engineering.

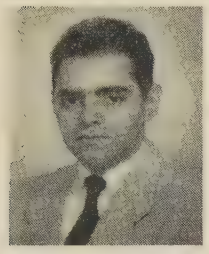
J. H. Richmond

In 1952 he received the degree of M.Sc. in electrical engineering at Ohio State University, and in 1955 the Ph.D. degree. Since 1952, Dr. Richmond has been engaged in research on circularly polarized antennas, rotary waveguide phase shifters, and radomes at the Antenna Laboratory of the Ohio State University. He has also been an assistant professor in the Department of Electrical Engineering since September, 1955.

Dr. Richmond is a member of Tau Beta Pi, Phi Beta Kappa, Sigma Xi, Pi Mu Epsilon, and Eta Kappa Nu.

❖

N. G. Sakiotis (S'48–A'50) was born in Lakeland, Fla., on March 5, 1928. He received the B.E.E. degree at the College of the City of New York in 1950.

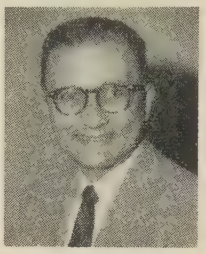


N. G. Sakiotis

After working with the Glenn L. Martin Co. and the Marine Radar Design Branch of the Bureau of Ships, Mr. Sakiotis joined the Microwave Antennas and Components Branch of the Naval Research Laboratory in 1951. While with the Laboratory, he has investigated figure-of-revolution scanners and is at present engaged in a study of the microwave properties of the ferrite materials.

❖

L. M. Spetner was born in St. Louis, Mo. on January 17, 1927. He received a B.S.M.E. degree from Washington University in 1945.



L. M. Spetner

After service in the U. S. Navy he returned to Washington University in 1946 and received an appointment as lecturer in the Department of Applied Mechanics until 1947. During this time he also worked as a research engineer on missile guidance and control problems with the Washington University Research Foundation.

From 1947 to 1950 he took graduate work in physics at Massachusetts Institute of Technology, doing research in cosmic rays and high energy particles. He received the Ph.D. degree from M.I.T. in 1950. During the summer of 1948 he worked in the servomechanism group at Emerson Electric Co. in St. Louis. From 1949 to 1950 he was a research assistant at M.I.T. and for one semester after receiving his degree he was a research associate.

From 1951 to the present Dr. Spetner has been a physicist with the Applied Physics Laboratory of the Johns Hopkins University where he has been engaged in research in electromagnetic propagation and noise theory.

He is a member of the American Physical Society, an associate member of Sigma Xi, and a member of Tau Beta Pi.

Michio Suzuki was born in Sapporo, Japan, on November 14, 1923. He received the B.S. degree in electrical engineering from Hakkaido University in Japan.

Michio Suzuki

Since 1948, Mr. Suzuki has been an assistant professor in the college of engineering at Hokkaido University.

He holds membership in the Institute of Electrical Communication Engineers of Japan.

❖

C. T. Tai (S'44–A'48–SM'51) was born in Soochow, China, in 1915. He received his B.S. degree from Tsing Hua University in 1937, and his D.Sc. degree from Harvard University in 1947. From 1947 to 1949 he was a research fellow of electronics at Harvard University. From 1949 to 1954 he was a research engineer at Stanford Research Institute. Since 1954 he has been an associate professor of electrical engineering at the Ohio State University, and an associate supervisor of the Antenna Laboratory.

C. T. Tai

❖

For a photograph and biography of J. R. Wait, see page 97 of the January, 1956 issue of Transactions of the PGAP.

Index to

IRE TRANSACTIONS

ON

ANTENNAS AND PROPAGATION

Volume AP-3, 1955

Reprinted from the January, 1956 issue for the benefit of those who wish to bind this index in with their 1955 volume.

IRE Transactions on Antennas and Propagation

Index to Volume AP-3—1955

Contents

Index to Authors

Numbers refer to index numbers in contents listing.

Index to Technical Subjects

Nontechnical Index